Applied Survival Analysis
Lab 10: Analysis of multiple failures

We will analyze the bladder data set (Wei et al., 1989). A listing of the dataset is given below:

```
list if id in 1/9

     +---------------------------------------------------------+
     | id   group   futime   number   size   r1   r2   r3   r4 |
     |---------------------------------------------------------|
  1. |  1       1        1        1      3    0    0    0    0 |
  2. |  2       1        4        2      1    0    0    0    0 |
  3. |  3       1        7        1      1    0    0    0    0 |
  4. |  4       1       10        5      1    0    0    0    0 |
  5. |  5       1       10        4      1    6    0    0    0 |
     |---------------------------------------------------------|
  6. |  6       1       14        1      1    0    0    0    0 |
  7. |  7       1       18        1      1    0    0    0    0 |
  8. |  8       1       18        1      3    5    0    0    0 |
  9. |  9       1       18        1      1   12   16    0    0 |
 10. | 10       1       23        3      3    0    0    0    0 |
     |---------------------------------------------------------|
 11. | 11       1       23        1      3   10   15    0    0 |
 12. | 12       1       23        1      1    3   16   23    0 |
 13. | 13       1       23        3      1    3    9   21    0 |
 14. | 14       1       24        2      3    7   10   16   24 |
 15. | 15       1       25        1      1    3   15   25    0 |
     |---------------------------------------------------------|
 16. | 16       1       26        1      2    0    0    0    0 |
 17. | 17       1       26        8      1    1    0    0    0 |
 18. | 18       1       26        1      4    2   26    0    0 |
 19. | 19       1       28        1      2   25    0    0    0 |
 20. | 20       1       29        1      4    0    0    0    0 |
     |---------------------------------------------------------|
 21. | 21       1       29        1      2    0    0    0    0 |
 22. | 22       1       29        4      1    0    0    0    0 |
 23. | 23       1       30        1      6   28   30    0    0 |
 24. | 24       1       30        1      5    2   17   22    0 |
 25. | 25       1       30        2      1    3    6    8   12 |
     +---------------------------------------------------------+
```

The data set is from a study in bladder cancer. The patients were followed for up to four recurrences (r1-r4). Some had less than four and some had none at all.

There are four ways to analyze these data that we will show below. These are:

- The Andersen-Gill (conditional model)
- The marginal (Wei-Lin-Weisfeld or WLW model)
- The conditional Prentice-Williams-Peterson (PWP) model. This has two verions:
  - The time from start model
  - The gap-time model

All of these models have in common that they attempt to describe the risk set (i.e., which subjects are at risk for which type of failure, first, second, third or fourth) and estimating the variance.

The Andersen-Gill model

This model (Andersen & Gill, 1981), assumes that the failures are ordered and each subject is at risk for failure *k* only after he or she has had failure *k*-1. That is, you cannot be at risk for the second failure before you have experienced the first failure. While this is a reasonable assumption, the model also assumes that the failures are *independent* from each other, that is, the model does not account for clustering of failures within the same subject.

The code to set up the A-G model is as follows:

```
. expand 5 if r4>0 & r4<futime
(48 observations created)

. expand 4 if !(r4>0 & r4<futime)
(219 observations created)

. sort id

. by id: gen rec=_n

. gen status=0

. gen tstart=0

. gen tstop=0

forvalues i=1/4 {
  2. replace status=1 if rec==`i' & r`i'>0 & r`i'<=futime
  3. replace tstop=r`i' if rec==`i' & r`i'>0 & r`i'<=futime
  4. replace tstart=tstop[_n-1] if rec==`i'& rec>1
  5. }

(47 real changes made)
(47 real changes made)
(0 real changes made)
(29 real changes made)
(29 real changes made)
(47 real changes made)
(22 real changes made)
(22 real changes made)
(29 real changes made)
(14 real changes made)
(14 real changes made)
(22 real changes made)

. by id:  replace tstart=tstop[_n-1] if rec==5
(12 real changes made)

. by id: drop if _n>1 &  tstart==0 & tstop==0
(157 observations deleted)

. by id: replace tstop=futime if _n==_N
(83 real changes made)

. drop if tstart==tstop
(5 observations deleted)

drop r1 r2 r3 r4
```

Here are two examples of subjects in the data (`id==9` and `id==25`)

```
. list if id==9 | id==25

     +-----------------------------------------------------------------------+
     | id   group   futime   number   size   rec   status   tstart   tstop |
     |-----------------------------------------------------------------------|
 11. |  9       1       18        1      1     1        1        0      12 |
 12. |  9       1       18        1      1     2        1       12      16 |
 13. |  9       1       18        1      1     3        0       16      18 |
 48. | 25       1       30        2      1     1        1        0       3 |
 49. | 25       1       30        2      1     2        1        3       6 |
     |-----------------------------------------------------------------------|
 50. | 25       1       30        2      1     3        1        6       8 |
 51. | 25       1       30        2      1     4        1        8      12 |
 52. | 25       1       30        2      1     5        0       12      30 |
     +-----------------------------------------------------------------------+
```

Subject 9 experienced two recurrences (at times 12 and 16) and was followed until time 18. That subject will have three observations with times 0-12, 12-16 and 16-18 and `status`=1 in the first two and `status=0` in the last observation. Similarly, subject 25 has experienced four recurrences up to time 12 and was followed up to time 30. That subject will have five entries with the latter censored.

The analysis is given as follows:

```
. stset tstop , fail(status) exit(time .) id(id) enter(tstart)

                id:  id
     failure event:  status != 0 & status < .
obs. time interval:  (tstop[_n-1], tstop]
 enter on or after:  time tstart
 exit on or before:  time .

------------------------------------------------------------------------------
      190  total obs.
        0  exclusions
------------------------------------------------------------------------------
      190  obs. remaining, representing
       85  subjects
      112  failures in multiple failure-per-subject data
     2711  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =        64
```

Note that we have to specify a starting time for each interval, otherwise STATA will consider each interval starting from time=0 (entry in the study). Given the A-G conditional assumption, this would have been incorrect since it would make each subject simultaneously eligible for all four failure types!

The analysis under the A-G model is given as follows:

```
. stcox group size number, nohr nolog

         failure _d:  status
   analysis time _t:  tstop
  enter on or after:  time tstart
  exit on or before:  time .
                 id:  id

Cox regression -- Breslow method for ties

No. of subjects =           85                     Number of obs   =       190
No. of failures =          112
Time at risk    =         2711
                                                   LR chi2(3)      =     14.05
Log likelihood  =   -460.07958                     Prob > chi2     =    0.0028

------------------------------------------------------------------------------
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       group |  -.4070966   .2000726    -2.03   0.042    -.7992317   -.0149615
        size |  -.0400877   .0702575    -0.57   0.568    -.1777899    .0976146
      number |   .1606478   .0480081     3.35   0.001     .0665536    .2547419
------------------------------------------------------------------------------
```

This analysis shows that the treatment `group` is protective of subsequent recurrences (HR=$e^{-0.4071} \approx 0.666$). On the other hand, the `number` of tumors prior to entry is related with the probability of subsequent recurrence (each additional tumor increases the risk of recurrence, on average, by 17% (HR=$e^{0.1606} \approx 1.174$).

The Wei-Lin-Weisfeld marginal model

The WLW model assumes that each tumor is a separate tumor type. Thus, the first tumor recurrence is a failure of type 1, the second of type 2 and so on. In addition, each subject is eligible for all recurrences (since they are simply failures of different types) *simultaneously*. While this is a mathematical approach (it is not logical in our setting of ordered failures) it makes sense in that, by setting the data in this manner, the approach allows construction of the correct matrices for calculation of the standard errors of the point estimates of the regression coefficients. The WLW approach uses a "sandwich estimator" of the variance of the type

$$V = I^{-1}G'GI^{-1} = D'D$$

where $I = \partial^2 \log L(\beta) / \partial\beta\partial\beta'$ is the usual information matrix and G is an $m \times p$ matrix of the score residuals. Matrix $D = GI^{-1}$ (is the matrix of leverage residuals – also called *dfbeta* by some packages) with elements d_{ij} that are the differences in the estimate of $\hat{\beta}_j$ if observation i is removed from the dataset. The WLW data set is constructed from the original bladder data set as follows:

```
. expand 4
(255 observations created)

. sort id
. by id: gen rec=_n
. gen status=0

. forvalues i=1/4 {
  2. replace status=1 if rec==`i' & r`i'>0 & r`i'<=futime
  3. replace futime=r`i' if rec==`i' & r`i'>0 & r`i'<=futime
  4. }

(47 real changes made)
(46 real changes made)
(29 real changes made)
(27 real changes made)
(22 real changes made)
(20 real changes made)
(14 real changes made)
(12 real changes made)

. drop r1 r2 r3 r4

. list if id <6

     +---------------------------------------------------+
     | id   group   futime   number   size   rec   status |
     |---------------------------------------------------|
  1. |  1       1        1        1      3     1        0 |
  2. |  1       1        1        1      3     2        0 |
  3. |  1       1        1        1      3     3        0 |
  4. |  1       1        1        1      3     4        0 |
  5. |  2       1        4        2      1     1        0 |
     |---------------------------------------------------|
  6. |  2       1        4        2      1     2        0 |
  7. |  2       1        4        2      1     3        0 |
  8. |  2       1        4        2      1     4        0 |
  9. |  3       1        7        1      1     1        0 |
 10. |  3       1        7        1      1     2        0 |
     |---------------------------------------------------|
 11. |  3       1        7        1      1     3        0 |
 12. |  3       1        7        1      1     4        0 |
 13. |  4       1       10        5      1     1        0 |
 14. |  4       1       10        5      1     2        0 |
 15. |  4       1       10        5      1     3        0 |
     |---------------------------------------------------|
 16. |  4       1       10        5      1     4        0 |
 17. |  5       1        6        4      1     1        1 |
 18. |  5       1       10        4      1     2        0 |
 19. |  5       1       10        4      1     3        0 |
 20. |  5       1       10        4      1     4        0 |
     +---------------------------------------------------+
```

The analysis of the WLW model with stata is as follows:

```
. stset futime, failure(status)

     failure event:  status != 0 & status < .
obs. time interval:  (0, futime]
 exit on or before:  failure

------------------------------------------------------------------------------
      340  total obs.
        0  exclusions
------------------------------------------------------------------------------
      340  obs. remaining, representing
      112  failures in single record/single failure data
     8522  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =        59
```

```
. stcox group size number, nohr efron strata(rec) cluster(id) nolog

         failure _d:  status
   analysis time _t:  futime

Stratified Cox regr. -- Efron method for ties

No. of subjects      =          340                Number of obs   =       340
No. of failures      =          112
Time at risk         =         8522
                                                   Wald chi2(3)    =     15.35
Log pseudolikelihood =   -426.14683                Prob > chi2     =    0.0015

                                    (Std. Err. adjusted for 85 clusters in id)
------------------------------------------------------------------------------
             |               Robust
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       group |  -.5847935   .3097738    -1.89   0.059    -1.191939    .0223521
        size |   -.051617    .095148    -0.54   0.587    -.2381036    .1348697
      number |   .2102937   .0670372     3.14   0.002     .0789032    .3416842
------------------------------------------------------------------------------
                                                              Stratified by rec
```

The main feature of the WLW model is that we account for the inter-subject clustering of the failures (i.e., repeated recurrences within the same subject cannot be assumed to be independent from each other), and that each failure is assumed to be its own stratum (i.e., different type of failure). These two features are addressed with the `strata(rec)` and `cluster(id)` options respectively.

The Prentice-Williams-Peterson model

There are two types of PWP models: The gap time model and the total time model. In both cases, the setup of the data set is identical to the A-G model, with the exception that time of observation past the last failure is not considered (i.e., once the fourth failure has occurred the patient is not considered further).

a) <u>The gap time model</u>
In this case, the PWP approach is a version of the A-G conditional model where each subject is considered at risk for each failure *conditional* on having experienced the previous failure. The differentiation of the model is in the fact that the variance estimation proceeds by a stratified analysis according to each failure (i.e., just as in the WLW model, the first failure is considered as failure of type 1, the second of type 2 and so on). In the gap-time model the length of the interval (i.e., (`tstart`, `tstop`]) is considered, where the start of the interval, just as in the A-G case, is past the occurrence of the previous failure (i.e., the subject cannot be eligible to experience a subsequent failure prior to having experienced all previous failures.

The setup of the data are similar to the A-G model, but the clock starts from the occurrence of the previous model. We will define variable `gap=tstop-tstart` and we will stset the data as follows:

```
. stset gap status

     failure event:  status != 0 & status < .
obs. time interval:  (0, gap]
 exit on or before:  failure

------------------------------------------------------------------------------
      183  total obs.
        5  obs. end on or before enter()
------------------------------------------------------------------------------
      178  obs. remaining, representing
      112  failures in single record/single failure data
     2480  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =        59
```

The analysis proceeds as in the case of single-observation per subject data, i.e., we do not include the `id()` option (that would produce an error by STATA)!

```
. stcox group size number, nohr nolog strata(rec)

         failure _d:  status
   analysis time _t:  gap

Stratified Cox regr. -- Breslow method for ties

No. of subjects =          178                     Number of obs   =       178
No. of failures =          112
Time at risk    =         2480
                                                   LR chi2(3)      =      8.76
Log likelihood  =   -363.16022                     Prob > chi2     =    0.0327

------------------------------------------------------------------------------
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       group |  -.2695213   .2076622    -1.30   0.194    -.6765318    .1374892
        size |   .0068402   .0700105     0.10   0.922    -.1303777    .1440582
      number |   .1535334   .0521059     2.95   0.003     .0514077     .255659
------------------------------------------------------------------------------
                                                             Stratified by rec
```

b) <u>The total time conditional model</u>

In this model, tstart is set to zero, i.e., the time at risk for each failure is the total time from entry until the occurrence of the failure.

The analysis of the PWP model proceeds as follows:

```
. stset tstop, fail(status) exit(time .) enter(t0)

     failure event:  status != 0 & status < .
obs. time interval:  (0, tstop]
 enter on or after:  time t0
 exit on or before:  time .

------------------------------------------------------------------------------
      183  total obs.
        0  exclusions
------------------------------------------------------------------------------
      183  obs. remaining, representing
      112  failures in single record/single failure data
     3907  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =        59
```

Note that we do not include the `id()` option (that would produce an error by STATA).

The analysis by the Cox model is given by the following output:

```
. stcox group size number, nohr nolog strata(rec)

         failure _d:  status
   analysis time _t:  tstop
  enter on or after:  time t0
  exit on or before:  time .

Stratified Cox regr. -- Breslow method for ties

No. of subjects =          183                     Number of obs   =       183
No. of failures =          112
Time at risk    =         3907
                                                   LR chi2(3)      =      8.75
Log likelihood  =   -367.17326                     Prob > chi2     =    0.0328

------------------------------------------------------------------------------
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       group |  -.4897246   .2092469    -2.34   0.019    -.8998411   -.0796082
        size |  -.0377304   .0675414    -0.56   0.576    -.1701092    .0946484
      number |   .1102692   .0510491     2.16   0.031     .0102149    .2103235
------------------------------------------------------------------------------
                                                              Stratified by rec
```