Nonparametric analysis: basic tools

Data description: The bladder cancer dataset contains recurrent event outcome information on 85 patients followed for the recurrence of bladder cancer tumor after transurethral surgical excision. The esposure of interest is the effect of the drug treatment of thiotepa. Control variables are initial number and intial size of tumors. The status variable is 1 for last recurrence and 0 for everything else (including death for any reason (only for practical purposes)). Time to event is recorded in months. The main aim of this study is to evaluate if the treatment is protective for recurrence.

1. Working on the pooled sample (85 patients):

```
cd "working directory path"
use bladder.dta,clear
```

2. Describe data looking at frequency distribution of variables by treatment and interpret the results

```
. describe

Contains data from bladder.dta
  obs:            85
 vars:             6
 size:           850 (99.9% of memory free)
-------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
id              byte   %8.0g                  identity number
rx              byte   %8.0g       type       treatment
number          byte   %8.0g                  initial number of tumours
size            byte   %8.0g                  size (cm) of largest initial tumour
stop            byte   %8.0g                  recurrence or censoring time
event           byte   %8.0g                  event of interest
-------------------------------------------------------------------------------
Sorted by:  id

. tabulate rx

  treatment |      Freq.     Percent        Cum.
------------+-----------------------------------
    placebo |         47       55.29       55.29
   thiotepa |         38       44.71      100.00
------------+-----------------------------------
      Total |         85      100.00

. tab number rx,col row chi2

   initial |
 number of |       treatment
```

```
   tumours |   placebo   thiotepa |     Total
-----------+----------------------+----------
         1 |        27         23 |        50
           |     54.00      46.00 |    100.00
           |     57.45      60.53 |     58.82
-----------+----------------------+----------
         2 |         8          3 |        11
           |     72.73      27.27 |    100.00
           |     17.02       7.89 |     12.94
-----------+----------------------+----------
         3 |         6          4 |        10
           |     60.00      40.00 |    100.00
           |     12.77      10.53 |     11.76
-----------+----------------------+----------
         4 |         3          1 |         4
           |     75.00      25.00 |    100.00
           |      6.38       2.63 |      4.71
-----------+----------------------+----------
         5 |         2          3 |         5
           |     40.00      60.00 |    100.00
           |      4.26       7.89 |      5.88
-----------+----------------------+----------
         6 |         0          2 |         2
           |      0.00     100.00 |    100.00
           |      0.00       5.26 |      2.35
-----------+----------------------+----------
         8 |         1          2 |         3
           |     33.33      66.67 |    100.00
           |      2.13       5.26 |      3.53
-----------+----------------------+----------
     Total |        47         38 |        85
           |     55.29      44.71 |    100.00
           |    100.00     100.00 |    100.00

          Pearson chi2(6) =   5.6363   Pr = 0.465

. sum size if rx==1,d

            size (cm) of largest initial tumour
-------------------------------------------------------------
      Percentiles      Smallest
 1%           1              1
 5%           1              1
10%           1              1       Obs                  47
25%           1              1       Sum of Wgt.          47

50%           1                      Mean           2.085106
                     Largest         Std. Dev.      1.529845
75%           3              5
```

```
90%            4              6       Variance       2.340426
95%            6              6       Skewness        1.54954
99%            7              7       Kurtosis       4.839122

. sum size if rx==2,d

             size (cm) of largest initial tumour
-------------------------------------------------------------
      Percentiles      Smallest
 1%            1              1
 5%            1              1
10%            1              1       Obs                  38
25%            1              1       Sum of Wgt.          38

50%            1                      Mean           1.921053
                       Largest       Std. Dev.      1.343283
75%            3              4
90%            4              4       Variance        1.80441
95%            5              5       Skewness       1.297615
99%            6              6       Kurtosis       3.825852


. ttest size,by(rx)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
 placebo |      47    2.085106    .2231508    1.529845    1.635927    2.534286
thiotepa |      38    1.921053    .2179093    1.343283    1.479526    2.362579
---------+--------------------------------------------------------------------
combined |      85    2.011765    .1565508    1.443327    1.700446    2.323083
---------+--------------------------------------------------------------------
    diff |            .1640538    .3162505               -.4649558    .7930633
------------------------------------------------------------------------------
    diff = mean(placebo) - mean(thiotepa)                         t =   0.5187
Ho: diff = 0                                     degrees of freedom =       83

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.6973         Pr(|T| > |t|) = 0.6053          Pr(T > t) = 0.3027

. tab event rx,col row chi2

  event of |       treatment
  interest |   placebo   thiotepa |     Total
-----------+----------------------+----------
         0 |        38         33 |        71
           |     53.52      46.48 |    100.00
           |     80.85      86.84 |     83.53
```

```
-----------+----------------------+----------
         1 |         9          5 |        14
           |     64.29      35.71 |    100.00
           |     19.15      13.16 |     16.47
-----------+----------------------+----------
     Total |        47         38 |        85
           |     55.29      44.71 |    100.00
           |    100.00     100.00 |    100.00

          Pearson chi2(1) =   0.5482   Pr = 0.459
```

3. Declare data to be survival-data in order to take into account the follow-up and interpret the new generated variables

```
. stset stop, fail(event) id(id)

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

--------------------------------------------------------------------------------
       85  total obs.
        0  exclusions
--------------------------------------------------------------------------------
       85  obs. remaining, representing
       85  subjects
       14  failures in single failure-per-subject data
     2480  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =        59
```

```
. list _st _d _t _t0 if _n<=5

     +---------------------+
     | _st   _d   _t   _t0 |
     |---------------------|
  1. |   1    0    1     0 |
  2. |   1    0    1     0 |
  3. |   1    0    1     0 |
  4. |   1    0    4     0 |
  5. |   1    0    5     0 |
     +---------------------+
```

4. Describe survival data

```
. stdes
                                    |-------------- per subject --------------|
Category                    total        mean         min     median        max
```

```
------------------------------------------------------------------------------
no. of subjects               85
no. of records                85           1           1           1           1

(first) entry time                         0           0           0           0
(final) exit time                   29.17647           1          27          59

subjects with gap              0
time on gap if gap             0           .           .           .           .
time at risk                2480    29.17647           1          27          59

failures                      14    .1647059           0           0           1
------------------------------------------------------------------------------
```

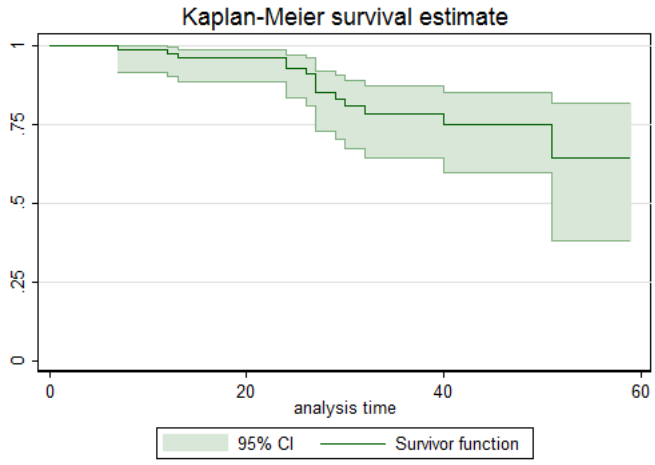
```
. stsum

         |                incidence       no. of    |------ Survival time -----|
         | time at risk      rate        subjects        25%       50%       75%
---------+--------------------------------------------------------------------
   total |         2480    .0056452            85         40         .         .
```

```
. stci

             |    no. of
             |  subjects         50%     Std. Err.     [95% Conf. Interval]
-------------+-------------------------------------------------------------
       total |        85           .             .           51           .
```

5. Estimate and plot the survival function by the Kaplan-Meier method with Greenwood confidence bands

```
. sts list

           Beg.          Net            Survivor      Std.
  Time    Total   Fail   Lost           Function     Error     [95% Conf. Int.]
-------------------------------------------------------------------------------
     1       85      0      3             1.0000         .          .         .
     4       82      0      1             1.0000         .          .         .
     5       81      0      1             1.0000         .          .         .
     7       80      1      1             0.9875    0.0124     0.9146    0.9982
     9       78      0      1             0.9875    0.0124     0.9146    0.9982
    10       77      0      3             0.9875    0.0124     0.9146    0.9982
    12       74      1      0             0.9742    0.0181     0.9005    0.9935
    13       73      1      1             0.9608    0.0222     0.8833    0.9872
---------
---------
    48       12      0      1             0.7482    0.0639     0.5964    0.8497
    49       11      0      2             0.7482    0.0639     0.5964    0.8497
```

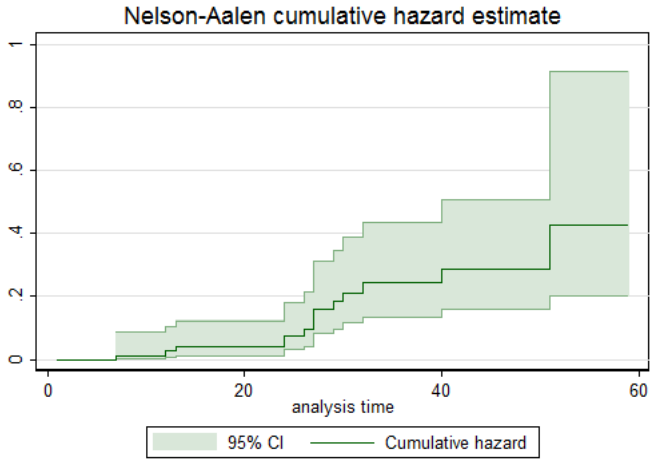
50	9	0	2	0.7482	0.0639	0.5964	0.8497
51	7	1	1	0.6413	0.1131	0.3801	0.8154
53	5	0	1	0.6413	0.1131	0.3801	0.8154
54	4	0	2	0.6413	0.1131	0.3801	0.8154
59	2	0	2	0.6413	0.1131	0.3801	0.8154



6. Estimate and plot the cumulative hazard function by the Aalen-Nelson with confidence bands

```
. sts list,na
```

Time	Beg. Total	Fail	Net Lost	Nelson-Aalen Cum. Haz.	Std. Error	[95% Conf.	Int.]
1	85	0	3	0.0000	0.0000	.	.
4	82	0	1	0.0000	0.0000	.	.
5	81	0	1	0.0000	0.0000	.	.
7	80	1	1	0.0125	0.0125	0.0018	0.0887
9	78	0	1	0.0125	0.0125	0.0018	0.0887
10	77	0	3	0.0125	0.0125	0.0018	0.0887
12	74	1	0	0.0260	0.0184	0.0065	0.1041
13	73	1	1	0.0397	0.0229	0.0128	0.1232
48	12	0	1	0.2847	0.0838	0.1599	0.5069
49	11	0	2	0.2847	0.0838	0.1599	0.5069
50	9	0	2	0.2847	0.0838	0.1599	0.5069
51	7	1	1	0.4276	0.1656	0.2001	0.9135
53	5	0	1	0.4276	0.1656	0.2001	0.9135
54	4	0	2	0.4276	0.1656	0.2001	0.9135
59	2	0	2	0.4276	0.1656	0.2001	0.9135



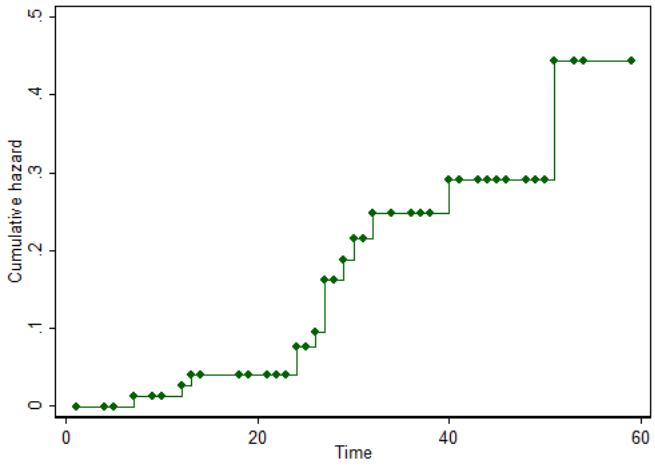
7. Provide an estimate for the cumulative hazard using the relationship between survival and hazard (S(t)=exp(-cumhaz(t))) and represent it

```
. sts generate km=s
. gen na_km=-log(km)
. sort stop
. list(stop na_km)
```

```
     +-----------------+
     | stop      na_km |
     |-----------------|
  1. |    1          0 |
  2. |    1          0 |
  3. |    1          0 |
  4. |    4          0 |
  5. |    5          0 |
     |-----------------|
  6. |    7   .0125788 |
  7. |    7   .0125788 |
  8. |    9   .0125788 |
---------
---------
 76. |   49   .2901115 |
 77. |   50   .2901115 |
 78. |   50   .2901115 |
 79. |   51   .4442622 |
 80. |   51   .4442622 |
     |-----------------|
 81. |   53   .4442622 |
 82. |   54   .4442622 |
 83. |   54   .4442622 |
 84. |   59   .4442622 |
 85. |   59   .4442622 |
```

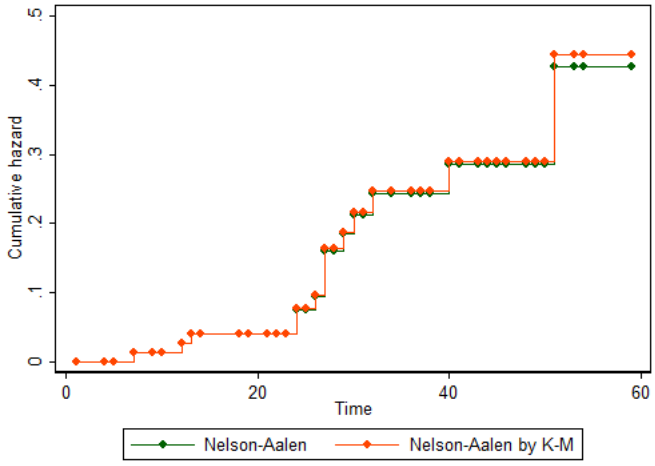
```
+-----------------+
```

```
. scatter na_km stop,connect(J) ytit("Cumulative hazard") xtit("Time")
```



8. Compare the results obtained in (e) and (f)

```
twoway (scatter na stop, connect(J)) (scatter na_km stop, connect(J)), ///
ytit("Cumulative hazard") xtit("Time") ///
legend (lab(1 "Nelson-Aalen") lab(2 "Nelson-Aalen by K-M"))
```



9. Estimate and plot the survival function by treatment

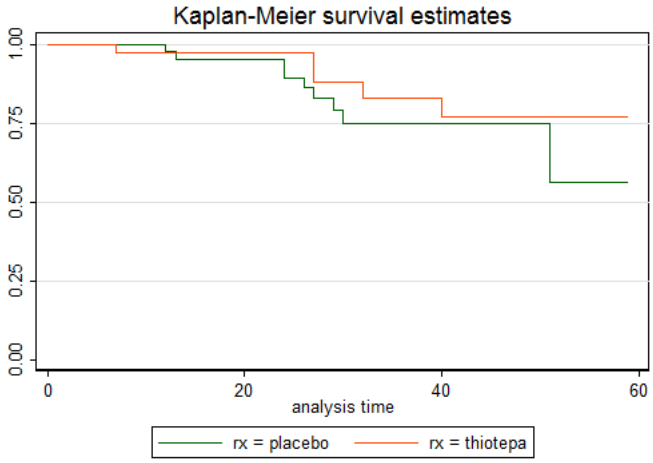
```
. sort rx
. sts list,by(rx)
```

```
          Beg.          Net            Survivor      Std.
Time      Total    Fail   Lost           Function      Error     [95% Conf. Int.]
```

```
-------------------------------------------------------------------------------
placebo
     1       47      0       1             1.0000        .          .          .
     4       46      0       1             1.0000        .          .          .
     7       45      0       1             1.0000        .          .          .
    10       44      0       2             1.0000        .          .          .
    12       42      1       0             0.9762   0.0235     0.8428     0.9966
    13       41      1       0             0.9524   0.0329     0.8227     0.9879
    14       40      0       1             0.9524   0.0329     0.8227     0.9879
    18       39      0       3             0.9524   0.0329     0.8227     0.9879
    23       36      0       4             0.9524   0.0329     0.8227     0.9879
    24       32      2       0             0.8929   0.0511     0.7372     0.9587
---------------------
---------------------
    48        6      0       1             0.7499   0.0789     0.5548     0.8689
    49        5      0       1             0.7499   0.0789     0.5548     0.8689
    51        4      1       1             0.5625   0.1728     0.1942     0.8170
    53        2      0       1             0.5625   0.1728     0.1942     0.8170
    59        1      0       1             0.5625   0.1728     0.1942     0.8170
thiotepa
     1       38      0       2             1.0000        .          .          .
     5       36      0       1             1.0000        .          .          .
     7       35      1       0             0.9714   0.0282     0.8140     0.9959
     9       34      0       1             0.9714   0.0282     0.8140     0.9959
    10       33      0       1             0.9714   0.0282     0.8140     0.9959
    13       32      0       1             0.9714   0.0282     0.8140     0.9959
    14       31      0       1             0.9714   0.0282     0.8140     0.9959
    18       30      0       2             0.9714   0.0282     0.8140     0.9959
---------------------
---------------------
    49        6      0       1             0.7681   0.0942     0.5190     0.8993
    50        5      0       2             0.7681   0.0942     0.5190     0.8993
    54        3      0       2             0.7681   0.0942     0.5190     0.8993
    59        1      0       1             0.7681   0.0942     0.5190     0.8993
-------------------------------------------------------------------------------

. sts graph,by(rx)
```

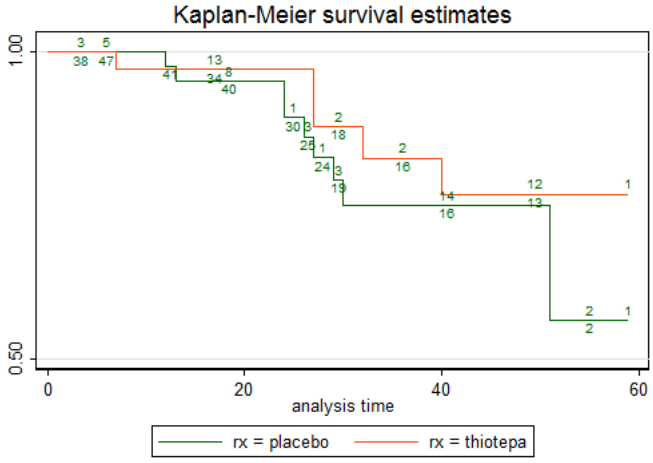
10. Compare the survival function at 19, 27, 41 months for the two treatment groups

```
. sts list, by(rx) compare at(19 27 41)

                 Survivor Function
rx              placebo   thiotepa
----------------------------------
time      19     0.9524     0.9714
          27     0.8276     0.8789
          41     0.7499     0.7681
----------------------------------
```



11. Represent the survival function by treatment showing also the number of patients at risk and censoring.

```
. sts graph, by(rx) lost atrisk ylab(0.50 1)
```



12. Estimate and plot the cumulative hazard function by treatment.

```
. sort rx
. sts list,by(rx) na

          Beg.          Net          Nelson-Aalen    Std.
  Time    Total    Fail   Lost          Cum. Haz.    Error     [95% Conf. Int.]
-------------------------------------------------------------------------------
placebo
     1       47       0      1             0.0000   0.0000          .          .
     4       46       0      1             0.0000   0.0000          .          .
     7       45       0      1             0.0000   0.0000          .          .
```

```
    10       44      0      2            0.0000    0.0000          .          .
    12       42      1      0            0.0238    0.0238     0.0034     0.1690
    13       41      1      0            0.0482    0.0341     0.0121     0.1927
    14       40      0      1            0.0482    0.0341     0.0121     0.1927
    18       39      0      3            0.0482    0.0341     0.0121     0.1927
--------------------
--------------------
    41       10      0      1            0.2813    0.1028     0.1374     0.5757
    43        9      0      2            0.2813    0.1028     0.1374     0.5757
    44        7      0      1            0.2813    0.1028     0.1374     0.5757
    48        6      0      1            0.2813    0.1028     0.1374     0.5757
    49        5      0      1            0.2813    0.1028     0.1374     0.5757
    51        4      1      1            0.5313    0.2703     0.1960     1.4401
    53        2      0      1            0.5313    0.2703     0.1960     1.4401
    59        1      0      1            0.5313    0.2703     0.1960     1.4401
thiotepa
     1       38      0      2            0.0000    0.0000          .          .
     5       36      0      1            0.0000    0.0000          .          .
     7       35      1      0            0.0286    0.0286     0.0040     0.2028
     9       34      0      1            0.0286    0.0286     0.0040     0.2028
    10       33      0      1            0.0286    0.0286     0.0040     0.2028
    13       32      0      1            0.0286    0.0286     0.0040     0.2028
    14       31      0      1            0.0286    0.0286     0.0040     0.2028
 --------------------
 --------------------
    45        9      0      1            0.2541    0.1180     0.1023     0.6312
    46        8      0      2            0.2541    0.1180     0.1023     0.6312
    49        6      0      1            0.2541    0.1180     0.1023     0.6312
    50        5      0      2            0.2541    0.1180     0.1023     0.6312
    54        3      0      2            0.2541    0.1180     0.1023     0.6312
    59        1      0      1            0.2541    0.1180     0.1023     0.6312
-------------------------------------------------------------------------------
```

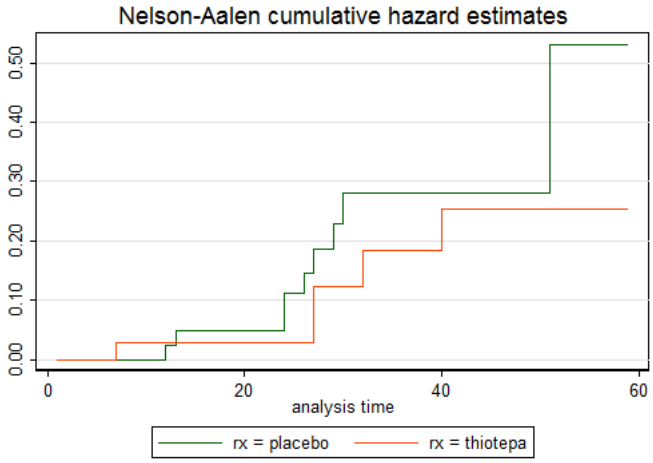
```
. sts graph,by(rx) na
```

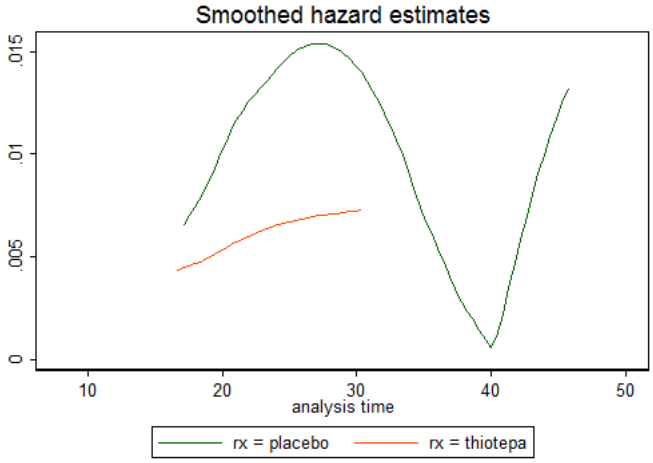
13. Represent the hazard function (smoothed estimate) by treatment

```
. sts graph,by(rx) hazard
```

14. Test the difference between the survival curves by treatment, by using i)log-rank test, ii)Wilcoxon-Breslow-Gehan test, iii) Tarone-Ware test, iv)Peto-Peto-Prentice test.

```
. sts test rx
Log-rank test for equality of survivor functions

         |   Events         Events
rx       |  observed       expected
---------+-------------------------
placebo  |         9           7.56
thiotepa |         5           6.44
```






```
---------+--------------------------
Total    |        14          14.00

               chi2(1) =       0.61
               Pr>chi2 =     0.4353

. sts test rx,wilcoxon
Wilcoxon (Breslow) test for equality of survivor functions

         |   Events         Events        Sum of
rx       |  observed       expected        ranks
---------+-------------------------------------
placebo  |         9           7.56           75
thiotepa |         5           6.44          -75
---------+-------------------------------------
Total    |        14          14.00            0
```

```
              chi2(1) =       0.62
              Pr>chi2 =     0.4315

. sts test rx,tware
Tarone-Ware test for equality of survivor functions

         |   Events         Events        Sum of
rx       |  observed       expected        ranks
---------+-------------------------------------
placebo  |         9           7.56     9.947012
thiotepa |         5           6.44    -9.947012
---------+-------------------------------------
Total    |        14          14.00            0

              chi2(1) =       0.61
              Pr>chi2 =     0.4362

. sts test rx,peto
Peto-Peto test for equality of survivor functions

         |   Events         Events        Sum of
rx       |  observed       expected        ranks
---------+-------------------------------------
placebo  |         9           7.56     1.304993
thiotepa |         5           6.44    -1.304993
---------+-------------------------------------
Total    |        14          14.00            0

              chi2(1) =       0.66
              Pr>chi2 =     0.4155
```

15. Estimate and plot the survival function by size of tumour

```
. tab size

  size (cm) |
 of largest |
    initial |
     tumour |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         48       56.47       56.47
          2 |         10       11.76       68.24
          3 |         16       18.82       87.06
          4 |          5        5.88       92.94
          5 |          2        2.35       95.29
          6 |          3        3.53       98.82
          7 |          1        1.18      100.00
------------+-----------------------------------
      Total |         85      100.00
```

```
. egen size_cat=cut(size),at(1,2,3,4,10)
. tab size_cat size
```

size_cat	size (cm) of largest initial tumour						
	1	2	3	4	5	6	7
1	48	0	0	0	0	0	0
2	0	10	0	0	0	0	0
3	0	0	16	0	0	0	0
4	0	0	0	5	2	3	1
Total	48	10	16	5	2	3	1

```
. sort size_cat
. sts list,by(size_cat)
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf.	Int.]
size_cat=1							
1	48	0	1	1.0000	.	.	.
4	47	0	1	1.0000	.	.	.
5	46	0	1	1.0000	.	.	.
7	45	0	1	1.0000	.	.	.
10	44	0	3	1.0000	.	.	.
12	41	1	0	0.9756	0.0241	0.8392	0.9965
13	40	0	1	0.9756	0.0241	0.8392	0.9965
14	39	0	1	0.9756	0.0241	0.8392	0.9965
18	38	0	3	0.9756	0.0241	0.8392	0.9965

51	4	1	1	0.5184	0.1659	0.1815	0.7765
54	2	0	1	0.5184	0.1659	0.1815	0.7765
59	1	0	1	0.5184	0.1659	0.1815	0.7765
size_cat=2							
9	10	0	1	1.0000	.	.	.
26	9	0	1	1.0000	.	.	.

41	3	0	2	0.8333	0.1521	0.2731	0.9747
45	1	0	1	0.8333	0.1521	0.2731	0.9747
size_cat=3							
1	16	0	2	1.0000	.	.	.
7	14	1	0	0.9286	0.0688	0.5908	0.9896
13	13	1	0	0.8571	0.0935	0.5394	0.9622

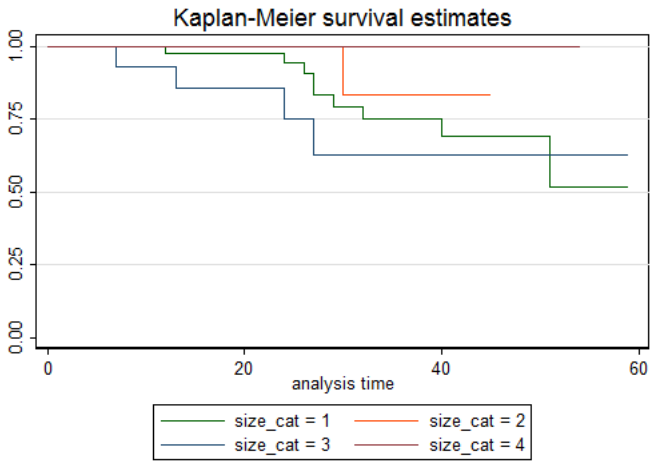
36	4	0	1	0.6250	0.1570	0.2619	0.8480

```
    49        3       0       2              0.6250     0.1570      0.2619      0.8480
    59        1       0       1              0.6250     0.1570      0.2619      0.8480
size_cat=4
    14       11       0       1              1.0000          .           .           .
    25       10       0       1              1.0000          .           .           .
    26        9       0       1              1.0000          .           .           .
---------
---------
    53        2       0       1              1.0000          .           .           .
    54        1       0       1              1.0000          .           .           .
-------------------------------------------------------------------------------------
```

```
. sts graph,by(size_cat)
. sts test size_cat
Log-rank test for equality of survivor functions
```

size_cat	Events observed	Events expected
1	9	7.50
2	1	2.00
3	4	1.97
4	0	2.53
Total	14	14.00

```
                chi2(3) =       5.51
                Pr>chi2 =     0.1381
```



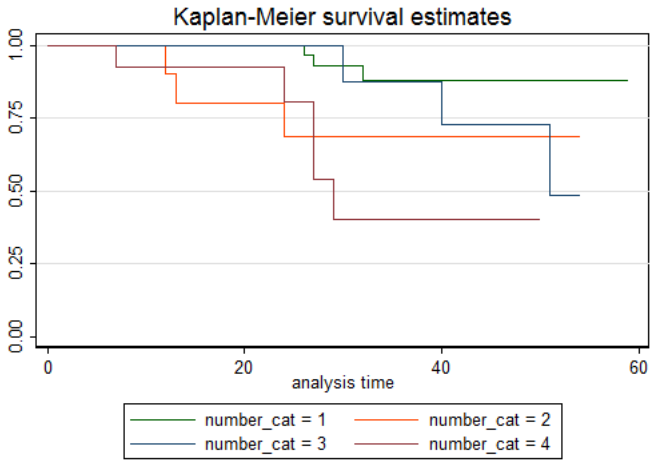
16. Estimate and plot the survival function by number of tumours

```
. tab number
```

```
    initial |
  number of |
    tumours |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         50       58.82       58.82
          2 |         11       12.94       71.76
          3 |         10       11.76       83.53
          4 |          4        4.71       88.24
          5 |          5        5.88       94.12
          6 |          2        2.35       96.47
          8 |          3        3.53      100.00
------------+-----------------------------------
      Total |         85      100.00

. egen number_cat=cut(number),at(1,2,3,4,10)
. tab number_cat number

           |                        initial number of tumours
number_cat |         1          2          3          4          5          6          8 |
-----------+-------------------------------------------------------------------------------+--
         1 |        50          0          0          0          0          0          0 |
         2 |         0         11          0          0          0          0          0 |
         3 |         0          0         10          0          0          0          0 |
         4 |         0          0          0          4          5          2          3 |
-----------+-------------------------------------------------------------------------------+--
     Total |        50         11         10          4          5          2          3 |


. sort number_cat
. sts list,by(number_cat)
           Beg.          Net            Survivor      Std.
  Time    Total   Fail   Lost            Function     Error     [95% Conf. Int.]
-------------------------------------------------------------------------------
number_cat=1
     1       50      0      3              1.0000         .          .          .
     7       47      0      1              1.0000         .          .          .
     9       46      0      1              1.0000         .          .          .
    10       45      0      1              1.0000         .          .          .
    13       44      0      1              1.0000         .          .          .
    14       43      0      1              1.0000         .          .          .
    18       42      0      4              1.0000         .          .          .
    21       38      0      1              1.0000         .          .          .
    22       37      0      1              1.0000         .          .          .
    23       36      0      2              1.0000         .          .          .
    25       34      0      4              1.0000         .          .          .
    26       30      1      3              0.9667    0.0328     0.7861     0.9952
    27       26      1      1              0.9295    0.0482     0.7457     0.9819
 -----------------
```

```
 ----------------
    50        4      0      1             0.8778    0.0677     0.6597    0.9600
    53        3      0      1             0.8778    0.0677     0.6597    0.9600
    59        2      0      2             0.8778    0.0677     0.6597    0.9600
number_cat=2
     4       11      0      1             1.0000         .          .         .
    12       10      1      0             0.9000    0.0949     0.4730    0.9853
    13        9      1      0             0.8000    0.1265     0.4087    0.9459
    14        8      0      1             0.8000    0.1265     0.4087    0.9459
    24        7      1      0             0.6857    0.1515     0.3046    0.8871
    29        6      0      1             0.6857    0.1515     0.3046    0.8871
    34        5      0      1             0.6857    0.1515     0.3046    0.8871
    36        4      0      1             0.6857    0.1515     0.3046    0.8871
    43        3      0      1             0.6857    0.1515     0.3046    0.8871
    44        2      0      1             0.6857    0.1515     0.3046    0.8871
    54        1      0      1             0.6857    0.1515     0.3046    0.8871
number_cat=3
    23       10      0      2             1.0000         .          .         .
    30        8      1      0             0.8750    0.1169     0.3870    0.9814
    36        7      0      1             0.8750    0.1169     0.3870    0.9814
    40        6      1      0             0.7292    0.1650     0.2764    0.9254
    41        5      0      1             0.7292    0.1650     0.2764    0.9254
    49        4      0      1             0.7292    0.1650     0.2764    0.9254
    51        3      1      1             0.4861    0.2269     0.0770    0.8163
    54        1      0      1             0.4861    0.2269     0.0770    0.8163
number_cat=4
     5       14      0      1             1.0000         .          .         .
     7       13      1      0             0.9231    0.0739     0.5664    0.9888
    10       12      0      2             0.9231    0.0739     0.5664    0.9888
    18       10      0      1             0.9231    0.0739     0.5664    0.9888
    19        9      0      1             0.9231    0.0739     0.5664    0.9888
    24        8      1      0             0.8077    0.1258     0.4098    0.9502
    26        7      0      1             0.8077    0.1258     0.4098    0.9502
    27        6      2      0             0.5385    0.1766     0.1740    0.8032
    29        4      1      1             0.4038    0.1765     0.0971    0.7029
    36        2      0      1             0.4038    0.1765     0.0971    0.7029
    50        1      0      1             0.4038    0.1765     0.0971    0.7029
-------------------------------------------------------------------------------

. sts graph,by(number_cat)
. sts test number_cat
Log-rank test for equality of survivor functions

           |   Events         Events
number_cat |  observed       expected
-----------+------------------------
1          |         3           7.98
2          |         3           1.85
3          |         3           2.67
```

```
4           |         5           1.50
------------+--------------------------
Total       |        14          14.00

                  chi2(3) =      12.29
                  Pr>chi2 =     0.0064
```



17. Estimate and plot the survival function by number of tumours and treatment

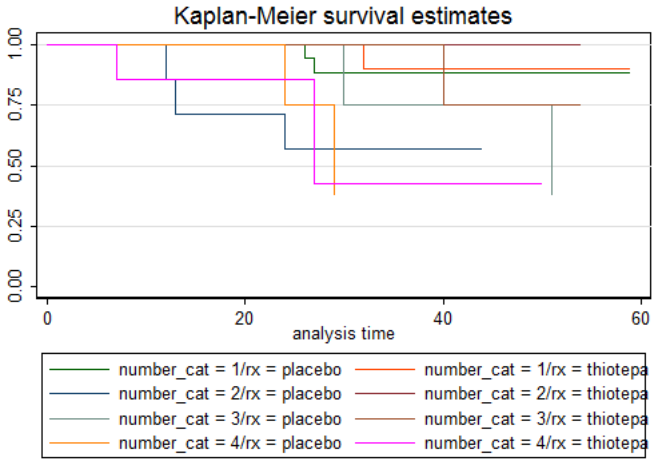
```
. tab number_cat rx

           |       treatment
number_cat |   placebo   thiotepa |     Total
-----------+----------------------+----------
         1 |        27          23 |        50
         2 |         8           3 |        11
         3 |         6           4 |        10
         4 |         6           8 |        14
-----------+----------------------+----------
     Total |        47          38 |        85
```

```
. sts graph,by(number_cat rx)
```

18. After re-loading data, estimate and plot the survival function by the Life-Table method on data grouped by year.

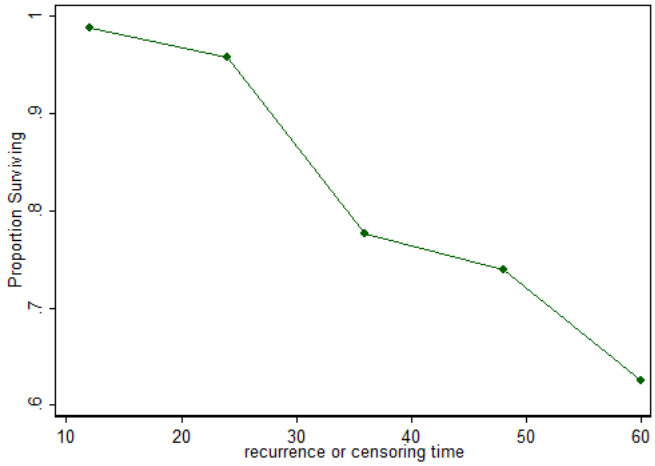
```
. use bladder.dta,clear
. ltable stop event,intervals(12)
```

```
                Beg.                                    Std.
   Interval     Total    Deaths    Lost    Survival     Error     [95% Conf. Int.]
-------------------------------------------------------------------------------
    0     12        85         1      10      0.9875    0.0124      0.9146     0.9982
```



12	24	74	2	15	0.9578	0.0239	0.8743	0.9863
24	36	57	9	19	0.7763	0.0578	0.6372	0.8674
36	48	29	1	16	0.7394	0.0658	0.5838	0.8441
48	60	12	1	11	0.6256	0.1185	0.3551	0.8086

```
. ltable stop event,intervals(12) graph
```

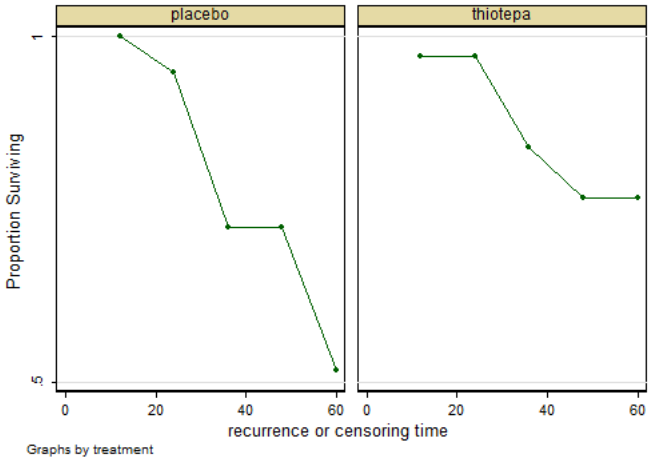


19. Estimate and plot the survival function by treatment by the Life-Table method on data grouped by year

```
. ltable stop event,intervals(12) graph by(rx)
```

Interval	Beg. Total	Deaths	Lost	Survival	Std. Error	[95% Conf.	Int.]
placebo							
0 12	47	0	5	1.0000	0.0000	.	.

```
   12     24         42          2          8       0.9474       0.0362       0.8056       0.9866
   24     36         32          6         13       0.7245       0.0843       0.5201       0.8531
   36     48         13          0          7       0.7245       0.0843       0.5201       0.8531
   48     60          6          1          5       0.5175       0.1850       0.1483       0.7966
thiotepa
    0     12         38          1          5       0.9718       0.0278       0.8164       0.9960
   12     24         32          0          7       0.9718       0.0278       0.8164       0.9960
   24     36         25          3          6       0.8393       0.0750       0.6211       0.9376
   36     48         16          1          9       0.7663       0.0978       0.5061       0.9012
   48     60          6          0          6       0.7663       0.0978       0.5061       0.9012
-------------------------------------------------------------------------------------------------
```



20. Estimate the mortality rate by the Life-Table method on data grouped by year

```
. ltable stop event,intervals(12) haz
```

```
                    Beg.      Cum.      Std.                  Std.
   Interval        Total   Failure     Error     Hazard      Error     [95% Conf. Int.]
-------------------------------------------------------------------------------------------------
    0     12         85    0.0125    0.0124     0.0010     0.0010     0.0000     0.0031
   12     24         74    0.0422    0.0239     0.0025     0.0018     0.0000     0.0061
   24     36         57    0.2237    0.0578     0.0174     0.0058     0.0061     0.0288
   36     48         29    0.2606    0.0658     0.0041     0.0041     0.0000     0.0120
   48     60         12    0.3744    0.1185     0.0139     0.0138     0.0000     0.0410
-------------------------------------------------------------------------------------------------
```

21. Load data bladder2.dta. It differs for previous data because information is given for both time at second last recurrence and time at last recurrence. The main aim of this study is to evaluate if the treatment is protective for recurrence conditional on having already suffered from a recurrence.

Working on the pooled sample (85 patients):

```
cd "working directory path"
use bladder2.dta,clear
```

22. Declare data to be survival-data in order to take into account the follow-up and interpret the new generated variables.

```
. stset stop, fail(event) entry(start) id(id)

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

------------------------------------------------------------------------------
       85  total obs.
        0  exclusions
------------------------------------------------------------------------------
       85  obs. remaining, representing
       85  subjects
       19  failures in single failure-per-subject data
     1659  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =        59

.
. list _st _d _t _t0 if _n<=5

     +----------------------+
     | _st   _d   _t   _t0 |
     |----------------------|
  1. |   1    0    1     0 |
  2. |   1    0    4     0 |
  3. |   1    0    7     0 |
  4. |   1    0   10     0 |
  5. |   1    0   10     6 |
     +----------------------+
```

23. Describe survival data.

```
. stdes

         failure _d:  event
   analysis time _t:  stop
  enter on or after:  time start
                 id:  id

                                    |-------------- per subject --------------|
Category                    total        mean         min     median        max
```

```
------------------------------------------------------------------------------
no. of subjects               85
no. of records                85           1           1           1           1

(first) entry time                  9.658824           0           2          47
(final) exit time                   29.17647           1          27          59

subjects with gap              0
time on gap if gap             0           .           .           .           .
time at risk                1659    19.51765           1          14          59

failures                      19    .2235294           0           0           1
------------------------------------------------------------------------------
```

```
. stsum

         |                incidence       no. of    |------ Survival time -----|
         | time at risk      rate        subjects        25%       50%       75%
---------+--------------------------------------------------------------------
   total |         1659    .0114527            85         27         .         .
```

24. Estimate and plot the survival function by the Kaplan-Meier method by treatment.

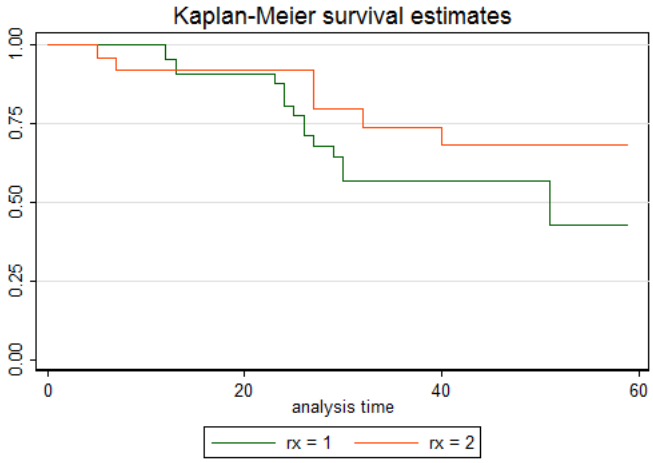
```
. sort rx

. sts list,by(rx)

        failure _d:  event
  analysis time _t:  stop
  enter on or after:  time start
                id:  id

           Beg.          Net            Survivor      Std.
  Time    Total   Fail   Lost           Function     Error     [95% Conf. Int.]
------------------------------------------------------------------------------
rx=1
     2       18      0     -1             1.0000         .          .         .
     3       19      0     -1             1.0000         .          .         .
     4       20      0      1             1.0000         .          .         .
     5       19      0     -1             1.0000         .          .         .
     6       20      0     -2             1.0000         .          .         .
     7       22      0      1             1.0000         .          .         .
     8       21      0     -1             1.0000         .          .         .
     9       22      0     -1             1.0000         .          .         .
    10       23      0      2             1.0000         .          .         .
    12       21      1     -1             0.9524    0.0465     0.7072    0.9932
    13       21      1      0             0.9070    0.0626     0.6769    0.9759
    14       20      0      1             0.9070    0.0626     0.6769    0.9759
    15       19      0     -2             0.9070    0.0626     0.6769    0.9759
```

```
   16       21       0      -3             0.9070    0.0626     0.6769    0.9759
 ---------------------
 ---------------------
   44        6       0       1             0.5673    0.0942     0.3655    0.7267
   46        5       0      -1             0.5673    0.0942     0.3655    0.7267
   48        6       0       1             0.5673    0.0942     0.3655    0.7267
   49        5       0       1             0.5673    0.0942     0.3655    0.7267
   51        4       1       1             0.4255    0.1417     0.1597    0.6716
   53        2       0       1             0.4255    0.1417     0.1597    0.6716
   59        1       0       1             0.4255    0.1417     0.1597    0.6716
rx=2
    1       21       0       2             1.0000         .          .         .
    2       19       0      -3             1.0000         .          .         .
    3       22       0      -1             1.0000         .          .         .
    5       23       1      -1             0.9565    0.0425     0.7293    0.9938
    6       23       0      -1             0.9565    0.0425     0.7293    0.9938
    7       24       1       0             0.9167    0.0564     0.7061    0.9785
    9       23       0       1             0.9167    0.0564     0.7061    0.9785
   10       22       0       1             0.9167    0.0564     0.7061    0.9785

  -------------------
  -------------------
   43       10       0       1             0.6810    0.1097     0.4173    0.8445
   44        9       0       1             0.6810    0.1097     0.4173    0.8445
   45        8       0       1             0.6810    0.1097     0.4173    0.8445
   46        7       0       2             0.6810    0.1097     0.4173    0.8445
   47        5       0      -1             0.6810    0.1097     0.4173    0.8445
   49        6       0       1             0.6810    0.1097     0.4173    0.8445
   50        5       0       2             0.6810    0.1097     0.4173    0.8445
   54        3       0       2             0.6810    0.1097     0.4173    0.8445
   59        1       0       1             0.6810    0.1097     0.4173    0.8445
-------------------------------------------------------------------------------
```

```
. sts graph,by(rx)
```

25. Test the difference between the survival curves by treatment, by using log-rank test.

```
. sts test rx
Log-rank test for equality of survivor functions

      |   Events         Events
rx    |  observed       expected
------+-------------------------
1     |        13          10.37
2     |         6           8.63
------+-------------------------
Total |        19          19.00

           chi2(1) =       1.51
           Pr>chi2 =     0.2191
```


Kaplan-Meier survival estimates
1.00
0.75
0.50
0.25
0.00
0
20
40
60
analysis time
rx = 1
rx = 2