Applied Survival Analysis
Solutions to lab 3: Comparing survival curves between groups

1. *Hemophiliac data set:*

The complete hemophiliac data set is given below. We have sorted it according to failure time to make our subsequent discussion easier.

```
. sort survival

. list, clean
```

```
        group   survival   censor
  1.      >40          1        1
  2.      >40          1        1
  3.      >40          1        1
  4.      >40          1        1
  5.     <=40          2        1
  6.      >40          2        1
  7.     <=40          3        0
  8.      >40          3        1
  9.      >40          3        1
 10.     <=40          6        1
 11.     <=40          6        1
 12.     <=40          7        1
 13.      >40          9        1
 14.     <=40         10        0
 15.     <=40         15        1
 16.     <=40         15        1
 17.     <=40         16        1
 18.      >40         22        1
 19.     <=40         27        1
 20.     <=40         30        1
 21.     <=40         32        1
```

So what is going to be the table in the first failure (*t*=1)?

***t*=1**

Group	Failure		Total
	Yes	No	
≤40	0	12	12
>40	4	5	9
Total	4	17	21

How about at time *t*=3? We must be careful here since there are only two failures and one censored observation, which will be removed from the "No" column after *t*=3, without adding a corresponding entry to the "Yes" column at *t*=3. That censored observation will be taken into account in the table associated with *t*=6. The table will be as follows:

***t*=3**

Group	Failure		Total
	Yes	No	
≤40	0	11	11
>40	2	2	4
Total	2	13	15

Finally, what will the table be like for *t*=10? This is a trick question! There will be no table entered for *t*=10 months, since there is only one censoring observation but no failures. This observation will be removed at *t*=15 months with the next two failures occurring in the younger than 40 group. The table at *t*=15 is as follows:

***t*=15**

Group	Failure Yes	Failure No	Total
≤40	2	4	6
>40	0	1	1
Total	2	5	7

Now, one way to analyze the data is to ask the question of whether there is an *association* between group membership and failure rates *adjusted across time*. This sounds like a Mantel-Haenszel statistic and indeed it is. The following Stata code involving the data set of the 2×2 tables is as follows:

```
. by t: tab group failure [weight=count]
```

```
-------------------------------------------------------------------------------
-> time = 1
(frequency weights assumed)

           |        failure
     group |        No        Yes |     Total
-----------+----------------------+----------
      <=40 |        12          0 |        12
       >40 |         5          4 |         9
-----------+----------------------+----------
     Total |        17          4 |        21


-------------------------------------------------------------------------------
-> time = 2

           |        failure
     group |        No        Yes |     Total
-----------+----------------------+----------
      <=40 |        11          1 |        12
       >40 |         4          1 |         5
-----------+----------------------+----------
     Total |        15          2 |        17


-------------------------------------------------------------------------------
-> time = 3

           |        failure
     group |        No        Yes |     Total
-----------+----------------------+----------
      <=40 |        11          0 |        11
       >40 |         2          2 |         4
-----------+----------------------+----------
     Total |        13          2 |        15


-------------------------------------------------------------------------------
-> time = 6

           |        failure
     group |        No        Yes |     Total
-----------+----------------------+----------
      <=40 |         8          2 |        10
       >40 |         2          0 |         2
-----------+----------------------+----------
     Total |        10          2 |        12
```

```
--------------------------------------------------------------------------------
-> time = 7

           |        failure
     group |        No        Yes |     Total
-----------+----------------------+----------
      <=40 |         7          1 |         8
       >40 |         2          0 |         2
-----------+----------------------+----------
     Total |         9          1 |        10


--------------------------------------------------------------------------------
-> time = 9

           |        failure
     group |        No        Yes |     Total
-----------+----------------------+----------
      <=40 |         7          0 |         7
       >40 |         1          1 |         2
-----------+----------------------+----------
     Total |         8          1 |         9


--------------------------------------------------------------------------------
-> time = 15

           |        failure
     group |        No        Yes |     Total
-----------+----------------------+----------
      <=40 |         4          2 |         6
       >40 |         1          0 |         1
-----------+----------------------+----------
     Total |         5          2 |         7


--------------------------------------------------------------------------------
-> time = 16

           |        failure
     group |        No        Yes |     Total
-----------+----------------------+----------
      <=40 |         3          1 |         4
       >40 |         1          0 |         1
-----------+----------------------+----------
     Total |         4          1 |         5


--------------------------------------------------------------------------------
-> time = 22

           |        failure
     group |        No        Yes |     Total
-----------+----------------------+----------
      <=40 |         3          0 |         3
       >40 |         0          1 |         1
-----------+----------------------+----------
     Total |         3          1 |         4


--------------------------------------------------------------------------------
-> time = 27

           |        failure
     group |        No        Yes |     Total
-----------+----------------------+----------
      <=40 |         2          1 |         3
-----------+----------------------+----------
     Total |         2          1 |         3
```

```
-----------------------------------------------------------------------------
-> time = 30

           |        failure
     group |        No        Yes |     Total
-----------+----------------------+----------
      <=40 |         1          1 |         2
-----------+----------------------+----------
     Total |         1          1 |         2


-----------------------------------------------------------------------------
-> time = 32

           |  failure
     group |       Yes |     Total
-----------+-----------+----------
      <=40 |         1 |         1
-----------+-----------+----------
     Total |         1 |         1
```

Now we perform the M-H analysis with STATA as follows:

```
. mhodds failure group [weight=count], by(t)
(frequency weights assumed)

Maximum likelihood estimate of the odds ratio
Comparing group==1 vs. group==0
by time

note: only 9 of the 12 strata formed in this analysis contribute
      information about the effect of the explanatory variable

-----------------------------------------------------------------------------
     time | Odds Ratio        chi2(1)         P>chi2       [95% Conf. Interval]
----------+------------------------------------------------------------------
        1 |          .           6.27         0.0122              .         .
        2 |   2.750000           0.44         0.5093        0.11974   63.15512
        3 |          .           5.92         0.0149              .         .
        6 |   0.000000           0.44         0.5071              .         .
        7 |   0.000000           0.25         0.6171              .         .
        9 |          .           3.50         0.0614              .         .
       15 |   0.000000           0.40         0.5271              .         .
       16 |   0.000000           0.25         0.6171              .         .
       22 |          .           3.00         0.0833              .         .
       27 |          .              .              .              .         .
       30 |          .              .              .              .         .
       32 |          .              .              .              .         .
-----------------------------------------------------------------------------

    Mantel-Haenszel estimate controlling for time
    ----------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ----------------------------------------------------------------
       4.725361       8.02        0.0046         1.443404  15.469704
    ----------------------------------------------------------------

Test of homogeneity of ORs (approx): chi2(8)   =   10.41
                                     Pr>chi2   =  0.2377
```

The p value associated with the M-H analysis is 0.0046 indicating that there is a significant association between group membership (i.e., age) and survival.

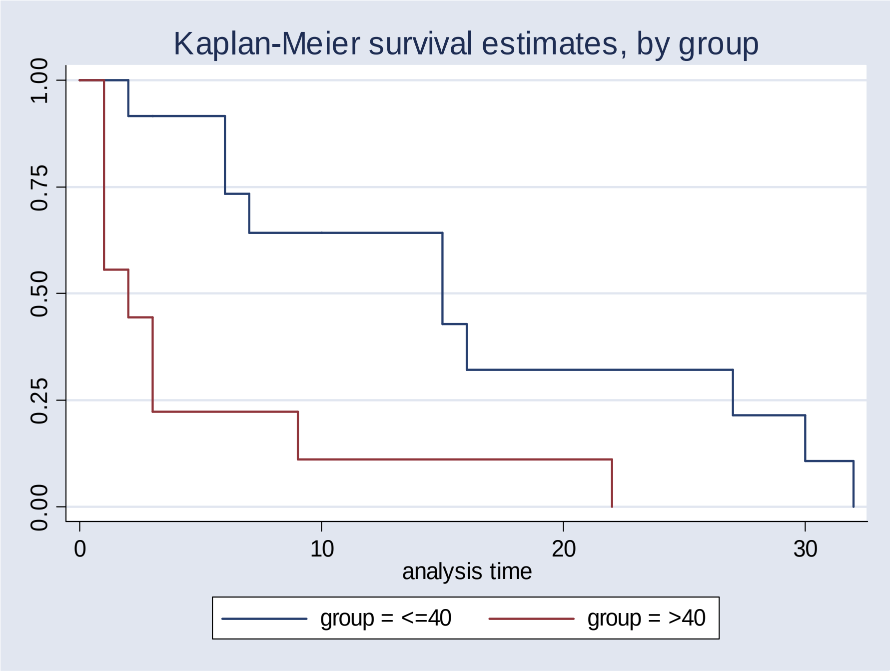
The statistical test thus constructed is called the *log-rank test*. It is calculate within command **`sts test group`** in Stata as follows:

```
. sts test group

         failure _d:  censor
   analysis time _t:  survival


Log-rank test for equality of survivor functions

      |   Events         Events
group |  observed       expected
------+------------------------
<=40  |        10          14.67
>40   |         9           4.33
------+------------------------
Total |        19          19.00

            chi2(1) =       8.02
            Pr>chi2 =     0.0046
```

The p value associated with the log-rank test is 0.0046 which is identical to the M-H analysis above. To see which of the two groups has the survival advantage, we can inspect the output and compare the median survival times. These are 2 months for the older group versus 15 months for the younger group.

Alternatively we can inspect the graph:

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



Since the survival curve associated with the younger than 40 group is consistently above the one associated with the older group, we conclude that the former enjoys a significant survival advantage compared to the latter.

2. *Leukemia Data:*

(a)

```
stset weeks remiss

     failure event:  remiss ~= 0 & remiss ~= .
obs. time interval:  (0, weeks]
 exit on or before:  failure

------------------------------------------------------------------------------
       42  total obs.
        0  exclusions
------------------------------------------------------------------------------
       42  obs. remaining, representing
       30  failures in single record/single failure data
      541  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =        35
```

(b)

```
sts graph, by(trt) l1(Survival Probability) b2(Time from Remission to
Relapse(weeks)) title(Comparison of Treatments for Leukemia)
```

The experimental group (6-MP) seems to be doing better than the control group. The relapse free curve is higher in the experimental group than in the control group.

(c) In both tests we would reject the null hypothesis of equality of the survival curves, since the p-values are highly significant ($p<0.0001$ and $p=0.0002$ and less than 0.05). So we would conclude that the survival curves are significantly different between the treatment groups (in favor of the experimental group). The Wilcoxon test puts more emphasis on early times and in this case the difference between the survival estimates in the beginning is not as big as in later times. So the Wilcoxon test statistic will be smaller and it's corresponding p-value will be larger (less significant) than the log-rank ($p=0.0002$ versus $p<0.0001$).