Applied Survival Analysis
Lab 3: Comparing survival curves between groups

In this lab, we are going to understand how to produce a Kaplan-Meier plot of survival estimates for more than one subgroup on the same graph, and compute a logrank or Wilcoxon test. Then we are going to see how to perform a stratified analysis in STATA.

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	
$\leq$40			
>40			
Total			

How about at time t=3? We must be careful here since there are only two failures and one censored observation, which is removed from the "No" column, but without adding a corresponding entry to the "Yes" column. The table will be as follows:

***t*=3**

Group	Failure		Total
	Yes	No	
$\leq$40			
>40			
Total			

Finally, what will the table be like for *t*=10

***t*=**

Group	Failure Yes	Failure No	Total
≤40			
>40			
Total			

Analyze the data is to ask the question of whether there is an *association* between group membership and failure rates *adjusted across time* using a Mantel-Haenszel statistic involving the data set of the 2×2 tables.

Now perform the log-rank test of the previous data set and compare it to the M-H analysis.

2. *Leukemia Data:*

First we are going to work with data from a leukemia remission study (Garrett 1997). The data consists of 42 patients who are followed over time to see how long (**`weeks`**) before they go out of remission (**`remiss`** : 1=yes, 0=no). Half of the patients received a new experimental drug and the other half received a standard drug (**`trt`** : 1=6-MP, 0=Control). This dataset is called *leukem.dta* Before we start the analysis we would like to label the values in the **`trt`** variable:

`label define trtfmt 1 "6-MP"  0 "Control"` (Define the format in a labelname **`trtfmt`**)
`label values trt trtfmt` (Assign the format to **`trt`**)

(a) Use the command discussed previously, to declare the data as survival type data.

If we would like the survival tables for the two treatment groups, we would type the usual **`sts list`** command but we would add the option **`by(covariate)`**, in this case **`trt`**. If the covariate has 2 levels (as here) , then we will get a 2-sample comparison. If the covariate has more levels, we will get a P-sample comparison. Note it must be a discrete covariate. If we add in the option **`compare`** you can get the survival functions of the two treatments side-by-side; it is usually easier to compare this way.

`sts list, by(trt)`

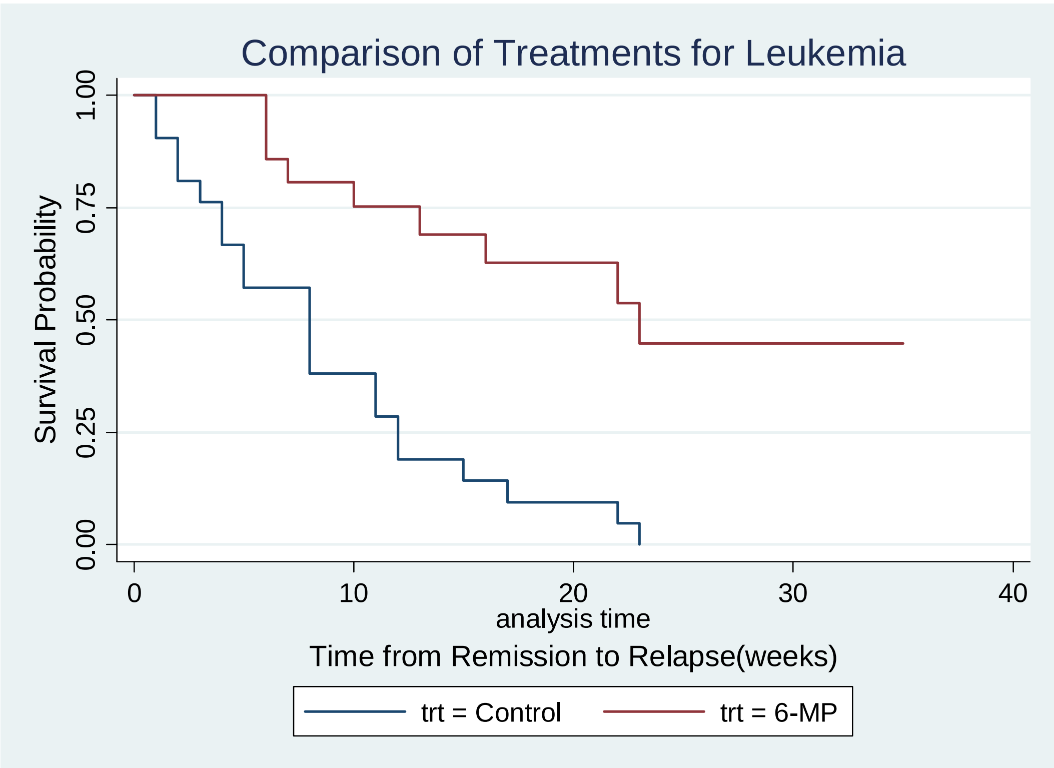
```
 failure _d:  remiss
 analysis time _t:  weeks

           Beg.          Net            Survivor      Std.
  Time    Total   Fail   Lost           Function     Error      [95% Conf. Int.]
-------------------------------------------------------------------------------
Control
     1       21      2      0             0.9048    0.0641     0.6700    0.9753
     2       19      2      0             0.8095    0.0857     0.5689    0.9239
     3       17      1      0             0.7619    0.0929     0.5194    0.8933
     4       16      2      0             0.6667    0.1029     0.4254    0.8250
     5       14      2      0             0.5714    0.1080     0.3380    0.7492
```

8	12	4	0	0.3810	0.1060	0.1831	0.5778
11	8	2	0	0.2857	0.0986	0.1166	0.4818
12	6	2	0	0.1905	0.0857	0.0595	0.3774
15	4	1	0	0.1429	0.0764	0.0357	0.3212
17	3	1	0	0.0952	0.0641	0.0163	0.2612
22	2	1	0	0.0476	0.0465	0.0033	0.1970
23	1	1	0	0.0000	.	.	.
6-MP							
6	21	3	1	0.8571	0.0764	0.6197	0.9516
7	17	1	0	0.8067	0.0869	0.5631	0.9228
9	16	0	1	0.8067	0.0869	0.5631	0.9228
10	15	1	1	0.7529	0.0963	0.5032	0.8894
11	13	0	1	0.7529	0.0963	0.5032	0.8894
13	12	1	0	0.6902	0.1068	0.4316	0.8491
16	11	1	0	0.6275	0.1141	0.3675	0.8049
17	10	0	1	0.6275	0.1141	0.3675	0.8049
19	9	0	1	0.6275	0.1141	0.3675	0.8049
20	8	0	1	0.6275	0.1141	0.3675	0.8049
22	7	1	0	0.5378	0.1282	0.2678	0.7468
23	6	1	0	0.4482	0.1346	0.1881	0.6801
25	5	0	1	0.4482	0.1346	0.1881	0.6801
32	4	0	2	0.4482	0.1346	0.1881	0.6801
34	2	0	1	0.4482	0.1346	0.1881	0.6801
35	1	0	1	0.4482	0.1346	0.1881	0.6801

Next we want to plot the survival functions of the two treatment groups:

(b) Guess what the command would be (it is similar to the above). What can you say about the two treatments based on the graph?



The general command in STATA to test for equality of survivor functions is **`sts test`** ***`covariate, options`*** . The default test is the *logrank* , adding the option **`wilcoxon`** or **`w`** give you the *Gehan's Wilcoxon Test*. An example of both tests follows:

`sts test trt` (Log-Rank Default)

```
         failure _d:  remiss
   analysis time _t:  weeks


Log-rank test for equality of survivor functions
------------------------------------------------

        |  Events
trt     |  observed       expected
--------+-------------------------
Control |        21          10.75
6-MP    |         9          19.25
--------+-------------------------
Total   |        30          30.00

              chi2(1) =      16.79
              Pr>chi2 =     0.0000
```

sts test trt, wilcoxon

```
         failure _d:  remiss
   analysis time _t:  weeks


Wilcoxon (Breslow) test for equality of survivor functions
----------------------------------------------------------

        |  Events                            Sum of
trt     |  observed       expected           ranks
--------+--------------------------------------
Control |        21          10.75             271
6-MP    |         9          19.25            -271
--------+--------------------------------------
Total   |        30          30.00               0

              chi2(1) =      13.46
              Pr>chi2 =     0.0002
```

(c) What would you conclude from the above results and why do you think the p-value of the Wilcoxon test is a bit higher (i.e., less significant) than the log-rank test ?