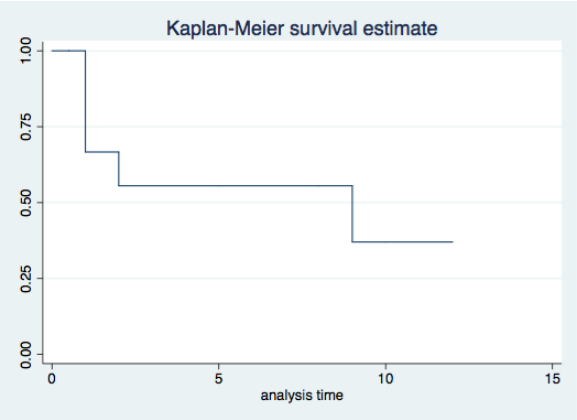
Exercise 10.1

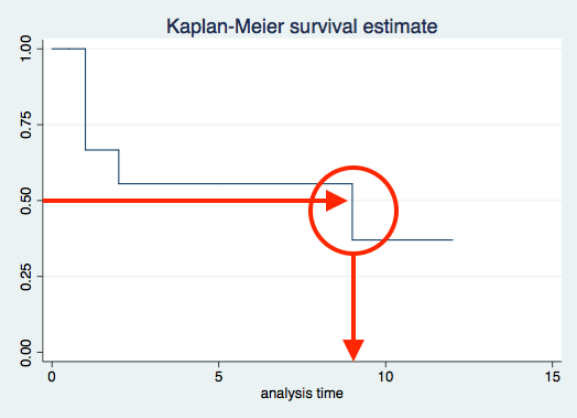
The survival times (in months) for 10 AIDS patients are as follows (*=censored):

0.5*,1,1,1,2,5*,8*,9,10*,12*

(a) A Kaplan-Meier curve was generated for these data using Stata. Use this to estimate the median survival time.



First time where the curve dips below 0.5 is just before 10 months; from the data we see this must be the 9 month data point – so the median survival time is 9 months.



The estimated survival function was obtained in Stata.

```
. sts list

         failure _d:  died
   analysis time _t:  months

           Beg.          Net            Survivor      Std.
  Time     Total    Fail   Lost           Function     Error     [95% Conf. Int.]
-------------------------------------------------------------------------------
    .5       10       0      1             1.0000         .          .         .
     1        9       3      0             0.6667    0.1571     0.2817    0.8783
     2        6       1      0             0.5556    0.1656     0.2042    0.8045
     5        5       0      1             0.5556    0.1656     0.2042    0.8045
     8        4       0      1             0.5556    0.1656     0.2042    0.8045
     9        3       1      0             0.3704    0.1872     0.0677    0.6933
    10        2       0      1             0.3704    0.1872     0.0677    0.6933
    12        1       0      1             0.3704    0.1872     0.0677    0.6933
-------------------------------------------------------------------------------
```

(b) What is the estimated survival probability at 1 month? At 5 months? At 6 months?

1 month: $\hat{S}(1) = 0.6667$
5 months: $\hat{S}(5) = 0.5556$
6 months: $\hat{S}(6) = 0.5556$

(c) Use the Stata output to confirm your answer to (a).

First value of $\hat{S}(t)$ *that is less than 0.5 is 0.3704, which first occurs at time* $= 9 \rightarrow$ *median survival is 9 months.*

Exercise 10.2

An observational study of mortality (time to death) for breast cancer patients was conducted. Patients were classified into two groups based on their tumor's immunohistochemical response (0=negative; 1=positive). Time to death or censoring was recorded in months.

The survival function estimate was separately for each group and a Kaplan-Meier plot created. A logrank test was also conducted. Stata output is on the next few pages. Use that output to answer the following questions.

(a) What is the estimated 25th percentile of survival for each group? Interpret these quantities.

Negative: 25th percentile = 56 months
→ We estimate that 25% of patients with negative tumors will survive less than 56 months.
Positive: 25th percentile = 38 months
→ We estimate that 25% of patients with positive tumors will survive less than 38 months.

(b) What is the estimated survival probability for each group at 2 years (24 months)? At 5 years (60 months)?

Negative: 97.22% survival at 2 years; 72.22% survival at 5 years
Positive: 77.78% survival at 2 years; 55.56% survival at 5 years

(c) What are the null and alternative hypotheses for the logrank test that was performed? Write down the hypotheses using symbols, and also explain in words what the null and alternative hypotheses mean.

Let $S_0(t)$ = survival function for negative group; $S_1(t)$ = survival function for the positive group
$H_0 : S_0(t) = S_1(t)$ → *the two survival functions are identical (same/overlap at all time points)*
$H_a : S_0(t) \neq S_1(t)$ → *they differ at at least one time point*

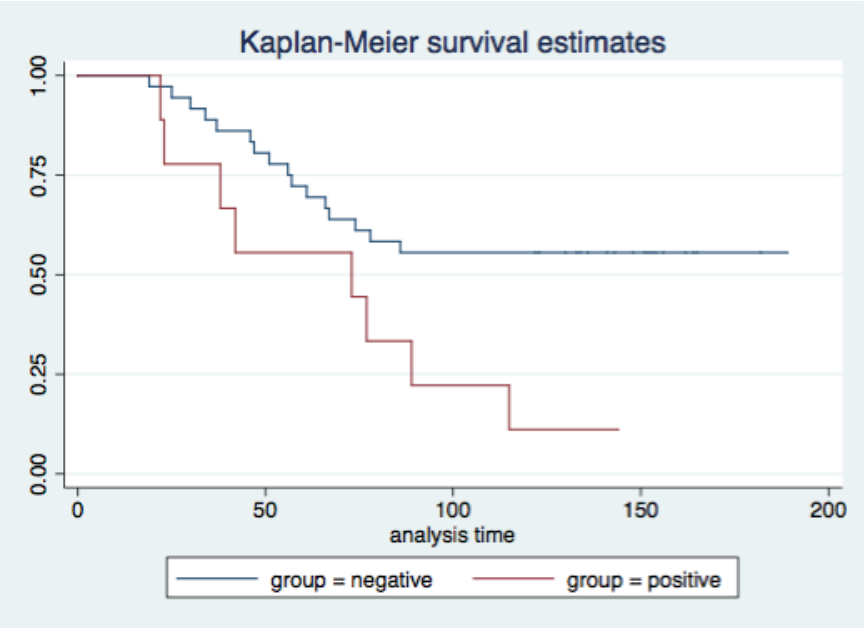
(d) Report the p-value for the logrank test and write a one-sentence conclusion (in context). (Assume $\alpha = 0.05$.)

p-value = 0.0191
There is evidence that survival (time to death) differs between patients with negative and positive tumors.

```
. sts list, by(group)

          Beg.          Net            Survivor      Std.
  Time    Total   Fail  Lost           Function      Error     [95% Conf. Int.]
-------------------------------------------------------------------------------
negative
    19       36      1     0             0.9722     0.0274     0.8187    0.9960
    25       35      1     0             0.9444     0.0382     0.7957    0.9858
    30       34      1     0             0.9167     0.0461     0.7635    0.9723
    34       33      1     0             0.8889     0.0524     0.7305    0.9568
    37       32      1     0             0.8611     0.0576     0.6980    0.9397
    46       31      1     0             0.8333     0.0621     0.6661    0.9214
    47       30      1     0             0.8056     0.0660     0.6350    0.9022
    51       29      1     0             0.7778     0.0693     0.6044    0.8821
    56       28      1     0             0.7500     0.0722     0.5746    0.8613
    57       27      1     0             0.7222     0.0747     0.5453    0.8398
    61       26      1     0             0.6944     0.0768     0.5165    0.8177
    66       25      1     0             0.6667     0.0786     0.4883    0.7950
    67       24      1     0             0.6389     0.0801     0.4607    0.7718
    74       23      1     0             0.6111     0.0812     0.4335    0.7482
    78       22      1     0             0.5833     0.0822     0.4067    0.7240
    86       21      1     0             0.5556     0.0828     0.3805    0.6994
   122       20      0     1             0.5556     0.0828     0.3805    0.6994
   123       19      0     1             0.5556     0.0828     0.3805    0.6994
```

```
   130        18       0       2              0.5556    0.0828     0.3805    0.6994
   133        16       0       1              0.5556    0.0828     0.3805    0.6994
   134        15       0       1              0.5556    0.0828     0.3805    0.6994
   136        14       0       1              0.5556    0.0828     0.3805    0.6994
   141        13       0       1              0.5556    0.0828     0.3805    0.6994
   143        12       0       1              0.5556    0.0828     0.3805    0.6994
   148        11       0       1              0.5556    0.0828     0.3805    0.6994
   151        10       0       1              0.5556    0.0828     0.3805    0.6994
   152         9       0       1              0.5556    0.0828     0.3805    0.6994
   153         8       0       1              0.5556    0.0828     0.3805    0.6994
   154         7       0       1              0.5556    0.0828     0.3805    0.6994
   156         6       0       1              0.5556    0.0828     0.3805    0.6994
   162         5       0       1              0.5556    0.0828     0.3805    0.6994
   164         4       0       1              0.5556    0.0828     0.3805    0.6994
   165         3       0       1              0.5556    0.0828     0.3805    0.6994
   182         2       0       1              0.5556    0.0828     0.3805    0.6994
   189         1       0       1              0.5556    0.0828     0.3805    0.6994
positive
    22         9       1       0              0.8889    0.1048     0.4330    0.9836
    23         8       1       0              0.7778    0.1386     0.3648    0.9393
    38         7       1       0              0.6667    0.1571     0.2817    0.8783
    42         6       1       0              0.5556    0.1656     0.2042    0.8045
    73         5       1       0              0.4444    0.1656     0.1359    0.7193
    77         4       1       0              0.3333    0.1571     0.0783    0.6226
    89         3       1       0              0.2222    0.1386     0.0337    0.5131
   115         2       1       0              0.1111    0.1048     0.0061    0.3877
   144         1       0       1              0.1111    0.1048     0.0061    0.3877
-------------------------------------------------------------------------------
```

(plot and more output for this problem on next page)

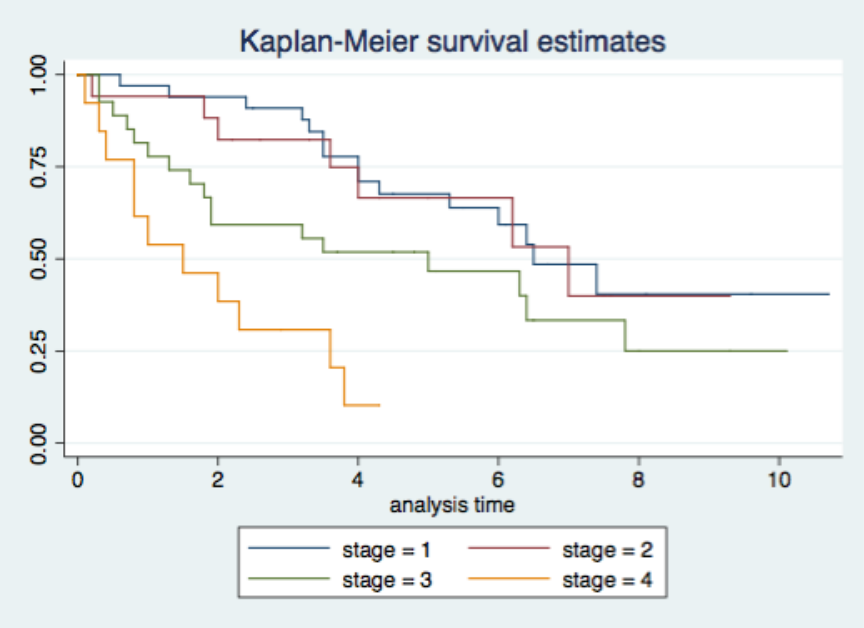



```
. sts test group, logrank

Log-rank test for equality of survivor functions

         |   Events         Events
group    |  observed       expected
---------+-------------------------
negative |        16          20.19
positive |         8           3.81
---------+-------------------------
Total    |        24          24.00

               chi2(1) =       5.49
               Pr>chi2 =     0.0191
```

Exercise 10.3

An observational study of survival of patients with cancer of the larynx was conducted. There were a total of 90 patients in the study, and for each patient we have information on the stage of cancer he/she has (stage = 1, 2, 3, 4). We also have the time in months until the patient died, was lost to follow-up, or was censored.

(a) Below is a Kaplan-Meier plot of these data. Describe what you see in terms of how survival might differ by cancer stage. (Which one has the "best" survival? Which one the "worst"?) Remember that you're not doing a test – just describing what you see.



In general, patients with higher cancer stage have worse survival – they die sooner. Patients with stage 4 cancer have the worst survival, followed by patients with stage 3 cancer. There doesn't appear to be much difference between stages 1 and 2, but both have better survival than stages 3 and 4.

(b) Using the plot, roughly approximate the median survival for each group.

Stage 1: 6.5 months; Stage 2: 7 months; Stage 3: 5 months; Stage 4: 1.5 months

(c) A logrank test was conducted; output is below. Write a one-sentence summary of the test result (in context), including reference to a p-value. (Assume $\alpha = 0.05$.)

```
. sts test stage, logrank

Log-rank test for equality of survivor functions

      |   Events         Events
stage |  observed       expected
------+-------------------------
1     |        15          22.57
2     |         7          10.01
3     |        17          14.08
4     |        11           3.34
------+-------------------------
Total |        50          50.00

            chi2(3) =      22.76
            Pr>chi2 =     0.0000
```

There is evidence that the survival (time to death) is not the same for all cancer stages ($p<0.00005$).

Note that we cannot definitively say which group(s) are significantly different from which other groups based on this test (would need additional tests – kind of like pairwise tests after an ANOVA).