Program Note 11.1 – Life Table calculations

The Stata command **ltable** can be used to perform life table calculations for survival data.

Program Note 11.2 – The Product Limit Method of Calculating Survival Probabilities

In Example 11.2, we present data on times to relapse of 14 alcohol-dependent patients shown in Table 11.4. We use several commands in Stata to summarize the time-to-relapse experience of 14 patients.

In the commands below, we demonstrate how data from Table 11.4 can be entered directly into Stata. The variable time refers to the time from discharge from a detoxification clinic to relapsing back to alcohol consumption. The variable relapse is an indicator that uses a "1" to identify those individuals who have relapsed, and the variable sex is used to differentiate between females and males where females = "1" and males = "2". Although we use "1" and "2" in this example, we could have generated a new variable gender coded as female = "0" and male = "1". The ordering of female and male is arbitrary, but we urge the reader to make it standard practice when applying numerical representations to nominal data to start with "0" as the base/reference category rather than "1".

```
Stata Command:

input time relapse sex
 4 1 2
 6 1 1
 6 1 2
 9 0 2
10 1 1
14 0 2
16 1 2
17 0 2
19 1 1
20 1 1
28 1 2
31 1 1
34 0 1
47 0 1
end
```

The command **stset** allows you to declare that the data is survival data. Following the **stset** command, the variable time is used to indicate the values of the event times. The variable time is followed by the **failure** statement which contains parentheses. Within the parentheses, we have indicated the censoring variable, relapse, as shown below:

```
Stata Command:

stset time, failure(relapse)

Stata Output:

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

------------------------------------------------------------------------------
      14  total obs.
       0  exclusions
------------------------------------------------------------------------------
      14  obs. remaining, representing
       9  failures in single record/single failure data
     261  total analysis time at risk, at risk from t =          0
                            earliest observed entry t =          0
                                last observed exit t =         47
```

After using the **stset** command to declare that the data are survival data, we can acquire Kaplan-Meier estimates of survival by using the **sts list** command.

```
Stata Command:

sts list

Stata Output:

        failure _d:  relapse
  analysis time _t:  time

          Beg.            Net      Survivor      Std.
  Time    Total   Fail   Lost     Function     Error      [95% Conf. Int.]
------------------------------------------------------------------------------
     4      14      1      0        0.9286     0.0688     0.5908     0.9896
     6      13      2      0        0.7857     0.1097     0.4725     0.9254
     9      11      0      1        0.7857     0.1097     0.4725     0.9254
    10      10      1      0        0.7071     0.1237     0.3938     0.8791
    14       9      0      1        0.7071     0.1237     0.3938     0.8791
    16       8      1      0        0.6188     0.1362     0.3075     0.8225
    17       7      0      1        0.6188     0.1362     0.3075     0.8225
    19       6      1      0        0.5156     0.1475     0.2136     0.7526
    20       5      1      0        0.4125     0.1497     0.1384     0.6727
    28       4      1      0        0.3094     0.1435     0.0784     0.5824
    31       3      1      0        0.2063     0.1274     0.0334     0.4804
    34       2      0      1        0.2063     0.1274     0.0334     0.4804
    47       1      0      1        0.2063     0.1274     0.0334     0.4804
------------------------------------------------------------------------------
```

The Stata command **sts graph** can be used to create Kaplan-Meier plots. The Stata commands are shown below. One brief note: we have added a vertical line to illustrate the median time occurring at 20 months.

```
sts graph, scheme(s1color)                              ///  fix the scheme
        xtitle({bf:Time in Months}, size(medlarge))      ///  modify the x-axis title
        ytitle({bf:Relapse Probablility}, size(medlarge)) ///  modify the y-axis title
        yscale(titlegap(4)) xscale(titlegap(4))          ///  creates title gaps
        title(" ")                                        ///  suppresses the main title
        xline(20, lpattern(dash) lcolor(gray))            ///  vertical dashed line
        plotopts(lcolor(black) lw(medthick))              ///  modify plot color and thickness
        risktable                                         ///  provides the number in risk set
                                                          ///  provides censoring indicator
        censored(single)
```

Stata Output:

Number at risk

| 14 | 10 | 5 | 3 | 1 | 0 |
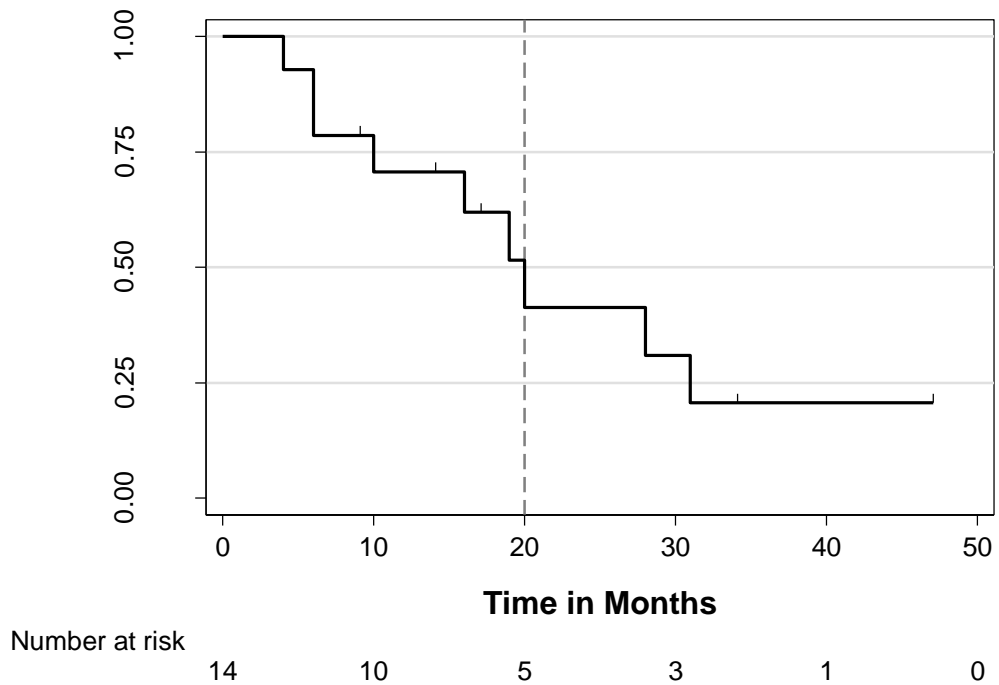
Program Note 11.3 – Comparing Survival Curves

In Example 11.4, we conducted the log-rank test to compare the survival function between males and females. The Stata command **sts test** followed by the variable sex— used to distinguish males from females— provides a log-rank test to assess if there is a difference in the survival experience between males and females.

```
Stata Commands:

sts test sex

Stata Output:

        failure _d:  relapse
   analysis time _t:  time

Log-rank test for equality of survivor functions

       |   Events           Events
sex    |  observed         expected
-------+---------------------------
1      |         5             6.19
2      |         4             2.81
-------+---------------------------
Total  |         9             9.00

          chi2(1) =       0.84
          Pr>chi2 =     0.3585




Stata Commands:

sts graph, scheme(s1color) title(" ")                     /// fix the scheme
        title(" ")                                        /// suppresses the main title
        by(sex)                                           /// the comparison groups
        xlab(0(12)48)                                     /// x-axis labels
        ylab(0(0.25)1.0)                                  /// y-axis labels
        censored(single)                                  /// the censoring indicator in the plot
        xtitle({bf:Time in Months})                       /// modify the x-axis title
        ytitle({bf: Relapse Probability})                 /// modify the y-axis title
        yscale(titlegap(4)) xscale(titlegap(4))           /// x-y scale options create title gaps
        ploto(lw(thick))                                  /// increase thickness of lines
        plot1(lcolor(gs7)  lpattern(dash))                /// line color and pattern for females
        plot2(lcolor(gs0) lpattern(solid))                /// line color and pattern for males
        risktable(, order(1  "Female:"                    /// number at risk
                       2  "Male:" ) size(small)) ///
        legend(pos(2) ring(0) region(lwidth(none)) size(small))     /// legend position
        legend(label(1 "Female") label(2 "Male") textfirst cols(1))  /// legend labels
        text(0.80 36.0 "Log-rank {&chi}{superscript:2}(1) = 0.84; P-value = 0.358", size(small))



Stata Output:
```
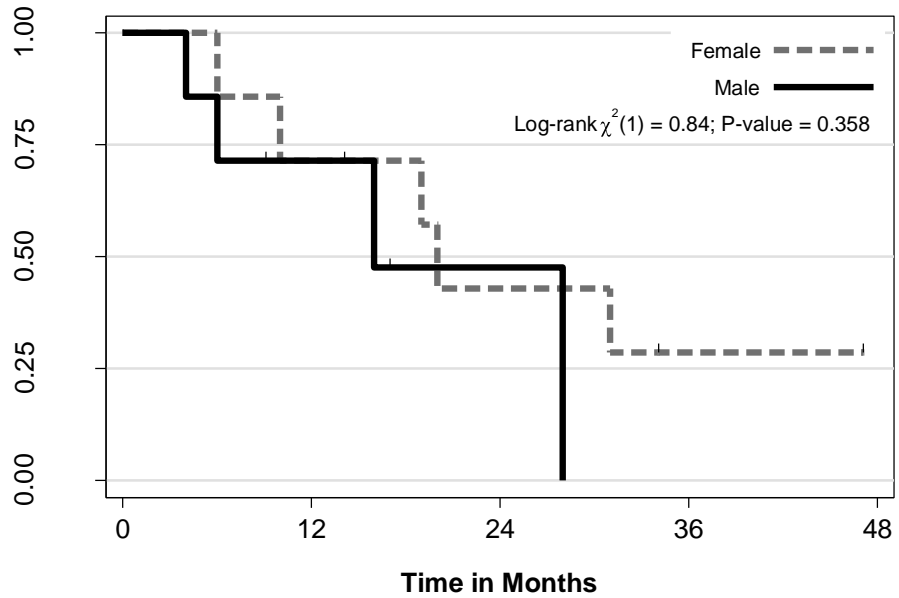
Female -- -- --
Male ———

Log-rank $\chi^2(1) = 0.84$; P-value = 0.358

Time in Months

Number at risk

Female:	7	5	3	1	0
Male:	7	4	1	0	0

Stata Commands:

```
* The median or 50th percentile for females (sex = 1)
stci if sex == 1, p(50)
```

Stata Output:

```
        failure _d:  relapse
  analysis time _t:  time

            |   no. of
            | subjects        50%     Std. Err.     [95% Conf. Interval]
------------+--------------------------------------------------------------
      total |        7         20     1.309307           6          .
```