Exercises: Module 1

1. **Open the WHAS500 data set in the software program of your choice.**

a. Produce a table of counts for fstat, to indicate which patients have died and which have been censored.

b. Draw a Kaplan-Meier plot for overall survival.

c. Estimate the 25th, 50th, and 75th quantiles for overall survival.

2. **Use the WHAS500 data set for this problem.**

a. Produce a crosstabulation of fstat and gender. Are you comfortable with the number of deaths in each group?

b. Draw Kaplan-Meier curves for males and females.

c. Calculate median survival with confidence intervals for males and females.

d. Calculate the log rank test for males versus females. Interpret your result.

3. **Use the WHAS500 data set for this problem.**

a. Produce age groups <60, 60-69, 70-79, and >=80. Compute a crosstabulation of this variable with fstat. Are you comfortable with the number of deaths in each group?

b. Draw Kaplan Meier curves for each age group.

c. Calculate the median survival time with confidence intervals for each age group.

d. Calculate the log rank test for age groups. Interpret your results.

4. **(Only for those who are brave) The following are times for catheters in infants. A "+" means that the catheter was removed because it was no longer needed. Times without a + mean that the catheter was removed because it failed. Occlusion and infection were the two major reasons for failure. Treating failures as an event and removal because it was no longer needed as a censored observation, estimate the Kaplan-Meier survival curve by hand, showing all your intermediate calculations.**

1+, 1+, 1+, 1+, 1+, 1+, 1+, 1+, 1, 1, 2+, 2+, 2, 2, 3+, 3, 4+, 4, 5+, 5+, 5+, 5+, 5+, 5, 5, 5, 6, 6, 7, 10, 10, 12, 12, 13