

ST745, Spring 2008

Homework 2 Due: Thursday, 02/7/2008

1. The following table shows data on time to HIV development for a sample of 100 individuals with STD but free of HIV at time 0:

Year intervals	# of HIV positive	# lost to follow-up
0-2	1	1
2-4	2	1
4-6	8	4
6-8	5	8
8-10	5	18
10-12	3	20
12-14	8	16

Use the data in this table to do the following (here we assume that censoring occurred in the middle of the interval):

- (a) Find the life-table estimate of the survival function of the time to HIV at years 6, 8, and 10 for the individuals with STD.
 - (b) Find the variance of the estimate you got in (a) at year 6, 8, and 10.
 - (c) Repeat the above using SAS and R.
2. For the following small data set of survival time: 3, 4, 5+, 6, 6+, 8, 11+, 14+, 15, 16+, where “+” means a right censored survival time, do the following:
- (a) Find the Kaplan-Meier estimate of the survival function and its variance at each failure time.
 - (b) Use the above Kaplan-Meier estimate to get an estimate and its variance of the cumulative hazard function at each failure time.
 - (c) Find the Nelson-Aalen estimate of the cumulative hazard function and its variance at each failure time.
 - (d) Find an estimate and its variance of the survival function using the Nelson-Aalen estimate you got in (c) at each failure time.

3. Using the lung cancer data (<http://www.biostat.mcw.edu/homepgs/klein/4.7.4.html>) in problem 4.3 of the textbook, do the following by using statistical software (such as SAS or R):
- (a) Find and plot the Kaplan-Meier estimate and the 90% pointwise confidence interval of the survival function using the data available on 3/31/1980.
 - (b) Find the estimate and its 90% CIs of the median survival time from the above plot using the method described in class. Compare your result to the output from SAS or R.