STAT 639V - Assignment 3

Carson Stacy

9/30/2021

1. Using the leukemia data (posted on Blackboard), compute the Nelson-Aalen estimators for the cumulative hazard function $H(t)$ and survival function $S(t)$ of 6-MP patients. (20 points)

The raw data used for this question is reproduced below. The estimators of this data are on the following pages.

Table 1: Remission duration of 6-MP vs placebo in children with acute leukemia

Pair	Remission Status at Randomization	Time to Relapse for Placebo Patients	Time to Relapse for 6-MP Patients
1	Partial Remissoin	1	10
2	Complete Remission	22	7
3	Complete Remission	3	32+
4	Complete Remission	12	23
5	Complete Remission	8	22
6	Partial Remissoin	17	6
7	Complete Remission	2	16
8	Complete Remission	11	34+
9	Complete Remission	8	32+
10	Complete Remission	12	25+
11	Complete Remission	2	11+
12	Partial Remissoin	5	20+
13	Complete Remission	4	19+
14	Complete Remission	15	6
15	Complete Remission	8	17+
16	Partial Remissoin	23	35+
17	Partial Remissoin	5	6
18	Complete Remission	11	13
19	Complete Remission	4	9+
20	Complete Remission	1	6+
21	Complete Remission	8	10+

The Nelson-Aalen estimates for Survival ($\tilde{S}(t)$) and Hazard ($\tilde{H}(t)$) functions for this data are calculated below.

t_i represents the i event times of the data, or times where at least one event occurred.

d_i represents the number of events that occurred at the i^{th} event time.

Y_i represents the number of subjects that were at risk for event occurring at time t_i.

$\tilde{H}$ represents the Nelson-Aalen estimate for hazard at time t_i and how it was calculated.

$\tilde{S}$ represents the Nelson-Aalen estimate for survival at time t_i and how it was calculated.

Table 2: Summary of Survival for Nelson-Aalen estimator

t_i	d_i	Y_i	$\tilde{H}$	$\tilde{S}$
6	3	21	3/21 = 0.1428571	exp(-0.1428571) = 0.8668779
7	1	17	(0.143)+(1/17) = 0.2016807	exp(-0.2016807) = 0.8173559
10	1	15	(0.202)+(1/15) = 0.2683473	exp(-0.2683473) = 0.7646422
13	1	12	(0.268)+(1/12) = 0.3516807	exp(-0.3516807) = 0.7035047
16	1	11	(0.352)+(1/11) = 0.4425898	exp(-0.4425898) = 0.6423707
22	1	7	(0.443)+(1/7) = 0.5854469	exp(-0.5854469) = 0.5568569
23	1	6	(0.585)+(1/6) = 0.7521136	exp(-0.7521136) = 0.4713692

From this data, we can say that NA estimator $\tilde{H}(t)$ for this data is:

$$\widetilde{H}(t) = \begin{cases} 0 & : t < 6 \\ 0.143 & : 6 \leq t < 7 \\ 0.202 & : 7 \leq t < 10 \\ 0.268 & : 10 \leq t < 13 \\ 0.352 & : 13 \leq t < 16 \\ 0.443 & : 16 \leq t < 22 \\ 0.585 & : 22 \leq t < 23 \\ 0.752 & : 23 \leq t < 35 \end{cases}$$

To find the standard deviation of the $\tilde{H}(t)$ for these time intervals:

$\sigma^2_H(t)$	$\sigma_H(t)$
3/(21*21) = 0.0068	sqrt(0.0068) = 0.0825
0.0068+(1/(17*17)) = 0.0103	sqrt(0.0103) = 0.1013
0.0103+(1/(15*15)) = 0.0147	sqrt(0.0147) = 0.1213
0.0147+(1/(12*12)) = 0.0217	sqrt(0.0217) = 0.1471
0.0217+(1/(11*11)) = 0.0299	sqrt(0.0299) = 0.1730
0.0299+(1/(7*7)) = 0.0503	sqrt(0.0503) = 0.2243
0.0503+(1/(6*6)) = 0.0781	sqrt(0.0781) = 0.2795

The NA estimate of the survival function $\tilde{S}(t)$ can also be determined from the information in Table 2 above:

$$\widetilde{S}(t) = \begin{cases} 1 & : t < 6 \\ 0.867 & : 6 \leq t < 7 \\ 0.817 & : 7 \leq t < 10 \\ 0.765 & : 10 \leq t < 13 \\ 0.704 & : 13 \leq t < 16 \\ 0.642 & : 16 \leq t < 22 \\ 0.557 & : 22 \leq t < 23 \\ 0.471 & : 23 \leq t < 35 \end{cases}$$

Therefore, the Nelson-Aalen estimators (estimate and standard deviations) for this data for both Hazard and Survival are summarized below:

$\tilde{H}(t)$	Time on Study (t)	Standard Deviation
0	$0 \leq t < 6$	0
0.143	$6 \leq t < 7$	0.0825
0.202	$7 \leq t < 10$	0.1013
0.268	$10 \leq t < 13$	0.1213
0.352	$13 \leq t < 16$	0.1471
0.443	$16 \leq t < 22$	0.1730
0.585	$22 \leq t < 23$	0.2243
0.752	$23 \leq t < 35$	0.2795

Time on Study (t)	$\tilde{S}(t)$	Standard Deviation
$0 \leq t < 6$	1	0
$6 \leq t < 7$	0.8669	0.0714
$7 \leq t < 10$	0.8174	0.0831
$10 \leq t < 13$	0.7646	0.0927
$13 \leq t < 16$	0.7035	0.1034
$16 \leq t < 22$	0.6424	0.1109
$22 \leq t < 23$	0.5569	0.1249
$23 \leq t < 35$	0.4714	0.1319