

Homework 2

Steven Chiou

Due date: Thursday, October 11

1. Show that (algebraically) in the absence of censoring $\hat{S}_{\text{KM}}(t) = \hat{S}_e(t)$.
2. In the absence of censoring, show that the Greenwood Formula (page 30 on note 2) can be reduced to

$$\frac{\hat{S}_{\text{KM}}(t) \times \{1 - \hat{S}_{\text{KM}}(t)\}}{n}.$$

You might assume there are no ties among the observations.

3. Consider the Leukemia data from the `survival` package:

```
> library(survival)
> head(aml)
```

	time	status	x
1	9	1	Maintained
2	13	1	Maintained
3	13	0	Maintained
4	18	1	Maintained
5	23	1	Maintained
6	28	0	Maintained

In here, each row represent one patient. `aml` is the observed survival time, `status` is the censoring indicator (1 = event, 0 = censored), and `x` is the treatment indicator. We will ignore the treatment indicator for now.

- a. Plot the Kaplan-Meier survival curve for the data.
 - b. Add the Nelson-Aalen survival curve to the Kaplan-Meier plot from (3a).
4. The expected survival time for the Leukemia data in #3 does not exist because the last observation is a censored event. An alternative is to look at the expected survival time, an alternative is to look at the restricted mean survival time. Compute $E(T|T < 161)$ based on the survival curve in (3a).
 5. Let $N_i(t)$ be the number of events over time interval $(0, t]$ for the i th patient in #3. Let $N(t) = \sum_{i=1}^n N_i(t)$ be the aggregated counting process.
 - a. Plot $N(t)$.
 - b. Plot $M(t)$, where $M(t) = N(t) - \hat{H}(t)$ and $\hat{H}(t)$ is the Nelson-Aalen estimator for the cumulative hazard function.
 - Note on 5b: After giving some thought, I think it is more meaningful to plot $dM(t) = dN(t) - \hat{h}(t)dt$. Both plots will receive full credit for 5b.