

P8110: Applied Regression II
Homework #2 [15 points]

NOTE: For the hand calculation questions, please only keep the first three decimals. Show all work.

1. Listed below are values of survival time in years for 10 patients, with the censored observations denoted by a "+" as a superscript: 0.4, 1.2, 1.2⁺, 3.4⁺, 4.3, 5.0, 5.0⁺, 6.1⁺, 7.1. The Kaplan-Meier estimates and 95% CIs of the survival functions are:

j	t_j	For t in	n_j	d_j	$1 - \frac{d_j}{n_j}$	$\hat{S}(t_j)$	95% CI
0	0.0	[0,0.4)	10	0	1	1	(1.0, 1.0)
1	0.4	[0.4,1.2)	10	1	0.9	0.9	(0.473, 0.985)
2	1.2	[1.2,4.3)	9	1	0.889	0.8	(0.409, 0.946)
3	4.3	[4.3,5.0)	6	1	0.833	0.667	(0.272, 0.882)
4	5.0	[5.0,7.1)	5	2	0.6	0.4	(0.097, 0.698)
5	7.1	[7.1,∞)	1	1	0	0	(0, 0)

- (a) Use SAS to re-generate the estimates of survival functions and 95% CIs in the table above. Show the SAS codes and outputs (only cut and paste the relevant SAS outputs) [2 points]
- (b) Hand calculate the median survival time and its 95% confidence interval. Interpret the results. [4 points]
- (c) Hand calculate the mean survival time estimate \hat{u}_1 and the variance estimate $\widehat{Var}(\hat{u}_1)$. [3 points]
- (d) Let \hat{u}_2 be the mean survival time estimate if all the censored observations were events. Without calculating $\hat{\mu}_2$, do you know which one of \hat{u}_1 and \hat{u}_2 is bigger? Briefly explain. Use SAS to estimate $\hat{\mu}_2$. [4 points]
- (e) Repeat the analyses in part (b) and (c) using SAS. Show the SAS codes and outputs (only cut and paste the relevant SAS outputs). [2 points]