P8110: Applied Regression II
Homework #3 [10 points]

NOTE: Please do not hand in computer output separately from responses to the questions below. Instead, cut and paste relevant SAS output. Attach SAS codes at the end of your homework.

The "HW3Data" gives the time until staphylococcus infection (in days) for 154 patients with a burn wound. The purpose of this study was to compare a routine bathing care treatment with a body-cleansing method. The time until staphylococcus infection was recorded, along with an indicator variable – whether or not an infection had occurred. The severity of the burn was measured by percentage of total surface area of body burned.

ID	=	Patient ID
treatment	=	0-routine bathing & 1-body cleansing
percentBurned	=	Percentage of total surface area burned
timeToInfection	=	Time to straphylocous aureaus infection or on study time
Infection	=	1-infection occured & 0-infection did not occur

1. Use SAS to compute the Kaplan-Meier estimator and 95% CI of the survival function for the routine bathing group (only keep the K-M estimates for the unique event time points). Generate a graph of the survival functions in the two treatment groups. Interpret the graph. (hint: What do you observe on the difference of the survival functions between the two groups?) [2 points]

2. Test whether the survival curves for the two treatment groups are the same using the log-rank test at $\alpha = 0.05$. Give the null and alternative hypothesis, test statistic, degrees of freedom, p-value, and conclusion. Does the Wilcoxon test lead to the same conclusion? Briefly explain why the test statistic of the Wilcoxon test is smaller than that of the log-rank test in this application. [5 points]

3. We can categorize the percentage of total surface area burned into four groups using the sample quartiles ($< 12\%, [12\%, 20\%), [20\%, 30\%), \text{and} \geq 30\%$). Test whether the survival functions are the same among the four groups using the generalized log-rank test at $\alpha = 0.05$. Give the null and alternative hypothesis, test statistic, degrees of freedom, p-value, and conclusion. [3 points]