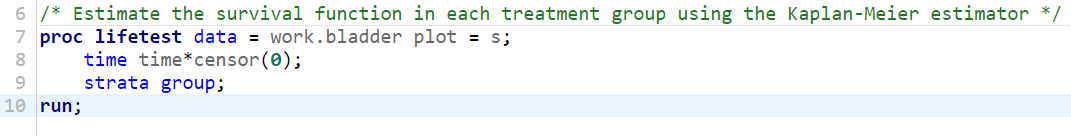
1. In a study of bladder cancer, tumors were removed from the bladders of 86 patients. Subsequently, the individuals were assigned to be treated either with the drug thiopeta or with a placebo. Time to the first recurrence of tumor in months is saved under the variable name time in the data set bladder.sas7bdat. Treatment status is saved under the variable group; the value 0 represents placebo. Indicators of censoring status - where 1 designates that a tumor did recur and 0 that it did not and that the observation was censored - are saved under variable censor. The variable number is an indicator of the number of tumors initially removed from the bladder: 1 indicates that a patient had a single tumor, and 2 that the individual had two or more tumors.
   1. Estimate the survival function in each treatment group using the Kaplan-Meier estimator. Don't need to include the whole table of Kaplan-Meier estimators here! Let S(t) = Pr(time > t) be the survival function of time, report Kaplan-Meier estimates of S(5), S(10) and S(25) for treatment group and placebo group, respectively. Given an explanation of Kaplan-Meier estimates of S(10) for treatment group and placebo group, respectively.



The LIFETEST Procedure

Stratum 1: group = 0

| Product-Limit Survival Estimates | | | | | |
| --- | --- | --- | --- | --- | --- |
| Time | Survival | Failure | Survival standard error | Number Failed | Number Left |
| 5 | 0.6947 | 0.3503 | 0.0681 | 14 | 31 |
| 10 | 0.5577 | 0.4423 | 0.0742 | 20 | 24 |
| 25 | 0.4040 | 0.5960 | 0.0760 | 26 | 14 |

For the placebo treatment group, the survival at S(10) is 0.5577, or 55.77%. The estimated chance of failure by 10 months is 0.4423, or 44.23%. There is a standard error of 0.0742. By this point in the study, 20 patients have failed, or had a tumor recur, and 24 have not experienced recurrence.

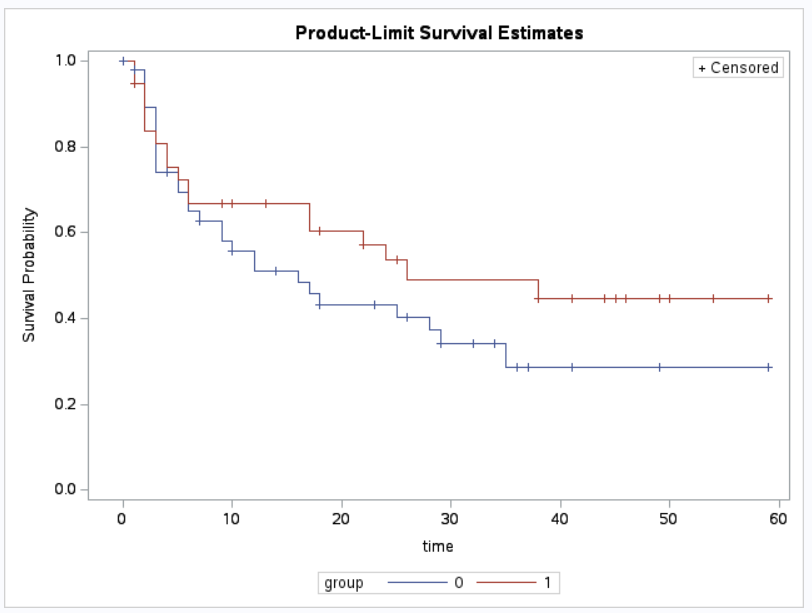
The LIFETEST Procedure

Stratum 2: group = 1

| Product-Limit Survival Estimates | | | | | |
| --- | --- | --- | --- | --- | --- |
| Time | Survival | Failure | Survival standard error | Number Failed | Number Left |
| 5 | 0.7245 | 0.2755 | 0.0743 | 10 | 26 |
| 10 | 0.6687 | 0.3313 | 0.0783 | 12 | 22 |
| 25 | 0.5357 | 0.4643 | 0.0867 | 16 | 14 |

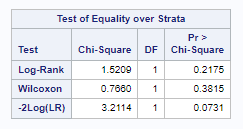
For the thiopeta treatment group, the survival at S(10) is 0.6687, or 66.87%. The estimated chance of failure by 10 months is 0.3313, or 33.13%. There is a standard error of 0.0783. By this point in the study, 12 patients have failed, or had a tumor recur, and 22 have not experienced recurrence.

* 1. Construct survival curves plot based on the Kaplan-Meier estimates. Does it appear that the individuals in one group have a longer time to first recurrence of tumor than those in the other group?



It appears based on the graph that both groups appear to experience recurrence at around 7-8 months into the study.

* 1. Test the null hypothesis that the distributions of recurrence times are identical in the two treatment groups. Which test do you use? What do you conclude?



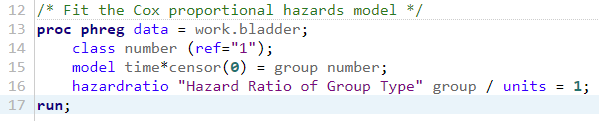
A log-rank test is a good test to use as it compares group versus survival status for each time t where a death occurs. Based on the output of the test of equality, the log-rank test has a chi-square value of 1.5209 for the difference of survival between the two groups. The p-value of 0.2175 is above the standard acceptance level of 0.05, which means we do not have sufficient evidence to reject the null hypothesis based on this log-rank test. This means that we cannot conclude that there is a statistically significant difference in survival time between the two treatment groups.

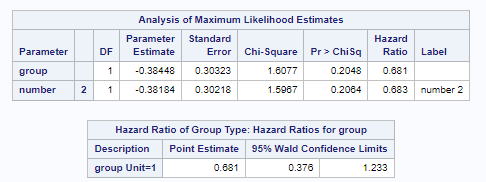
* 1. Write down the expression for Cox proportional hazards model where group and number are explanatory variables. Remember to define the hazard functions in the model.

h(t|X) = h0(t)exp(βTX) = h0(t)exp(β1x1i + … + βpxpi)

h(t|X) = h0(t)exp(β1treatment +β2number), where h(t|X) represents the hazard function at time t for the subject, h0(t) represents the baseline hazard function, β1treatment represents the variable coefficient and the treatment group value, and β2number represents the variable coefficient and number captures the initial number of tumors.

* 1. Fit the Cox proportional hazards model in part (d). Estimate the hazard ratio of the treatment group versus placebo group, and get its 95% confidence interval. Interpret the results.





The 95% confidence interval of the treatment group vs. placebo group is (0.376, 1.233). This ratio indicates that there is a lot of uncertainty in the model, as one part of the Wald interval is below 1, and the other is above 1. So while it may appear that the treatment is not effective, the uncertainty of the model demonstrates that additional samples and further research may be required to draw a definitive conclusion.