

Survival Analysis
Biostatistics 140.641
Due date: end of 3rd term, 2006

Problem Set 3 Obtain the data set from course website
<http://www.biostat.jhsph.edu/~xli/teaching/survival/survival.html>
Due day: End of 3rd term

Data descriptions: This data set (Crowley and Hu, 1977) consists of 103 cardiac patients who enrolled in the transplantation program at the Stanford University Hospital between 1967 and 1974. After enrollment, each patient waited a length of time and received a heart transplant surgery when a suitable donor heart was found. Thirty patients died before a donor heart was identified, and four patients stayed alive without a heart transplant at the termination date of the study (April 1, 1974). Of the 69 transplant recipients, only 24 were still alive at the end of the study. At the time of transplantation, all but four of the patients were tissue typed to determine the degree of similarity with the donor. The covariates of interest are list as follows:

DOB: date of birth.
DOA: date of acceptance into the study.
DOT: date of transplant.
DLS: date last seen (dead or censored).
DEAD: coded 1 if death occurs at DLS; coded 0 otherwise.
SURG: coded 1 if patient had open-heart surgery prior to DOA; coded 0 otherwise.
M1: number of donor alleles with no match in recipient (1 through 4).
M2: coded 1 if donor-recipient mismatch on HLA-A2 antigen, coded 0 otherwise.
M3: mismatch score.
AGEACCPT: age (in years) at acceptance into the program
AGETRANS: age at transplant
TRANS: coded 1 for those who received a transplant and coded 0 for those who did not.
SURV1: days from acceptance to death, and
SURV2: days from heart transplant to death.

WAIT: days from acceptance to heart transplant.

Remark: Please pay attention to the way to define the ‘censored data’. An observed survival/failure time should be treated as a censored observation if death did not occur prior to the date last seen.

1. Compare the failure time, SURV1, from patients with or without prior open-heart surgery prior to DOA. Answer this question by plotting the Kaplan-Meier estimates, and performing nonparametric test procedures for the two groups. Are you ‘comfortable’ with these statistical methods? Discuss your results.
2. Define the failure time as the time from acceptance to heart transplant (i.e., WAIT). Answer the questions stated in Problem 1.
3. Use SURV1 as the failure time and ($AGEACCP$ T, $SURG$) as the covariates. Use the proportional hazards model to conduct regression analysis. Discuss your results.
4. Use SURV1 as the failure time and ($AGEACCP$ T, $SURG$, $TRANS(t)$) as the covariates, where $TRANS(t)$ is a time-dependent transplant indicator. Use the proportional hazards model to conduct regression analysis. Discuss your results.
5. Compare the proportional hazards models used in Problems 3 and 4. Which model is more relevant for analysis?
6. Use SURV1 as the failure time and consider a two sample analysis to compare the group of patients with the heart transplants v.s. the group without. Compute and plot the Kaplan-Meier estimates. Is such a comparison relevant? If the answer is no, is there an alternative approach to use instead of the Kaplan-Meier estimates? Discuss your results.

*****IMPORTANT***** Do not turn in all the computing sheets. Just turn in the key computing outputs together with your discussion. Try to make your homework as concise as possible!