Homework No. 5

The first problem of this homework continues the use of the Worcester Heart Attack data which can be found in the Homework Assignments folder.

Define the survival time of a subject as the time between admission to the hospital and death. If a subject is still alive, the survival time is treated as censored.

1. Define a time-dependent covariate called cumulative hosiptal stay, i.e., the number of days the patient had stayed in the hosiptal at a given time point. Determine whether this time-dependent covariate alone has an effect on the survival of a patient.

2. Check the effect of the time-dependent covariate defined in the previous question after adjusted for the time-independent covariates `age, gender, hr, diasbp, bmi, sho, chf, year`. What is the conclusion you draw from the analysis?

3. In homework 4, you checked the proportional hazards assumption by comparing the baseline hazards for different strata. It appears that the proportional hazards assumption may not be satisfied. An alternative way to check is to include time-dependent covariates in the form time*covariate. Define such time-dependent covariates for each variable selected in the final model in homework 4 and then perform a variable selection with all these time-dependent covariates as well as the time-independent covariates. Write down the model you selected.

4. Interpret the model you selected in the previous question. What are the implications of your result? Based on the result of your analysis, is there any way to further check the proportional hazards assumption?