Homework No. 3

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. Compare the Kaplan-Meier estimator of the overall survival function with that of the exponential model fit in question 2 of Homework No. 2, the exponential model fit appears not appropriate. Fit a Weibull model to the survival data and summarize the fitted model.

2. Plot the Weibull model fit and the Kaplan-Meier estimator in a single graph and comment on the Weibull model fit.

3. Perform a formal test to see if the Weibull model fit significantly improves over the exponential model fit.

4. Perform a test to compare male and female survival functions based on Weibull model fits. What is your conclusion based on this test?

5. Divide the subjects into four groups based on their age ($< 60, 60-70, 70-80, \geq 80$) and then test to see if the survival functions for different age groups are the same or not based on the Weibull model fit, where age category $60-70$ means $60 \leq age < 70$.

6. Is the effect of `age` on the survival linear in the Weibull model?

7. Exam individually the effect of `hr`, `sysbp`, `diasbp`, and `bmi` on survival using the Weibull model.

8. Fit a Weibull model adjusting for possible effects of `gender`, `age`, `hr`, `sysbp`, `diasbp`, and `bmi`.

9. Perform a variable selection to exclude unimportant covariates from the model in the previous question and summarize your final model fit.

10. Interpret your final model fit.