Homework No. 4

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. Fit proportional hazards regression models with one covariate at a time for each of the covariates `age, gender, hr, sysbp, diasbp, bmi, cvd, afb, sho, chf, av3, miord, mitype, year`. Summarize your results in a single table.

2. Select from the list of covariates that are significant at type I error level 0.15 and fit a proportional hazards regression model with all the selected covariates.

3. Perform variable selection from the covariates used in the previous question.

4. Interpret the model parameter estimates in the model you fitted in the previous question.

5. Draw a graph to show your estimate of the survival function at the means of continuous variables and category 0 (except variable `year` which is at category 1) of discrete variables from the proportional hazards model fit in the previous question.

6. Perform a stratified proportional hazards regression analysis with stratification variables `gender`.

7. Compare the stratified proportional hazards regression model fit and the unstratified model fit (after the model selection). Discuss their similarity and difference.

8. Perform a separate proportional hazards regression analysis for each stratum defined by `gender`. Estimate the baseline cumulative hazards (i.e., negative logarithm of the survival function).

9. Plot of the log-log transformed baseline survival functions obtained in the previous problem against the logarithm of the survival time in a single graph. Comment on the proportional hazards assumption on the variable `gender` by visually check the graph.

10. Perform similar analyses in questions 6, 7, 8, 9 with the variable `age` categorized by $< 70, 70 - 80, > 80$ years old in place of `gender`. Comment on what you may learn from such analyses.