Final Project (Please type your report and hand in before **June 12, Monday**)

Consider the data from Prevention of Events with Angiotensin Converting Enzyme Inhibition (PEACE) Trial. The goal of the study was to test whether ACE-inhibitor therapy, when added to modern conventional therapy, would reduce the rate of nonfatal myocardial infarction, death from cardiovascular causes, or revascularization in low-risk patients with stable coronary artery disease and normal or slightly reduced left ventricular function. Patients underwent randomization from November 1996 to June 2000 and were followed up for as long as 7 years (median, 4.8 years), until December 31, 2003. The study was conducted after approval from the institutional review boards at 187 sites in the United States (including Puerto Rico), Canada, and Italy. Patients gave their written informed consent to participate. An independent data and safety monitoring board reviewed patient safety data and interim results. A morbidity and mortality review committee reviewed and classified all outcomes. The data consist of the following variables

```
t2death: time to death (months)
death: censoring status (1=death; 0=censored)
tx: 0=standard 1=treatment
age: age(years) at baseline
sysbp: systolic blood pressure at baseline
gender: 1=female; 0=male
hidiabet: history of diabetes (1=yes; 0=no) at baseline
hihypert:   history of hypertension (1=yes; 0=no) at baseline
```

The complete data set be read into R using

```
data=read.table("peacedata.csv", head=T, sep=",")
```

1. Conduct the logrank test to test the treatment effect of ACE-inhibitor therapy in reducing mortality

2. Estimate the hazard ratio of the ACE-inhibitor versus the standard care only and construct the associated confidence interval based on the Cox regression model. Report your findings. Compare the p-value of the treatment effect with that from the logrank test. Why are they almost identical?

3. It is known that age, systolic blood pressure, gender, history of diabetes and history of hypertension are associated with the survival time. Estimate the hazard ratio of the the ACE-inhibitor versus the standard care only but adjusting for the aforementioned factors, using the multivariate Cox regression model. Report your findings.

4. Estimate the hazard ratio of the ACE-inhibitor versus the standard care only and construct the associated confidence interval based on the Cox regression model in male and female patients, separately. Test if these two hazard ratios are identical. Report and interpret your findings.

5. The clinical investigator decides to develop prognostic regression models using the baseline age, systolic blood pressure, gender, history of diabetes and history of hypertension to predict the survival time for patients receiving the conventional therapy only and for patients receiving the ACE-inhibitor plus the conventional therapy. To this end, one may build two separate Cox regression models in patient receiving the conventional therapy only (tx=0) and in patient receiving the ACE-inhibitor plus the conventional therapy (tx=1).

   (a) Plot the estimated survival functions for following four patients:

   - patient A receiving the conventional therapy only (age=60, sysbp=140, gender=1, hidiabet=0, hihypert=1)
   - patient B receiving the ACE-inhibitor plus conventional therapy (age=140, sysbp=60, gender=1, hidiabet=0, hihypert=1)
   - patient C receiving the conventional therapy only (age=60, sysbp=140, gender=0, hidiabet=0, hihypert=1)
   - patient D receiving the ACE-inhibitor plus conventional therapy (age=140, sysbp=60, gender=0, hidiabet=0, hihypert=1)

   Would you give different treatment recommendations for a 60-year old male patient, who has a systolic blood pressure of 140 and history of hypertension but has no diabetes, and a female patient with the same characteristics? Why?

   (b) The researcher decides to use the restricted mean survival time (up to 80 months) to summarize the survival curve. What are the RMST for patients A and B based on your estimated survival curves.

   (c) You may use the resampling method to construct the 95% confidence interval for these two RMSTs (please see the reference paper as well as the lecture slides). The basic idea is to replace $dM_i(t)$ by $dN_i(t)G_i$, where $G_i \sim N(0,1)$ generated by the users. Describe your procedure and construct the corresponding 95% confidence intervals.