Homework 8

Angela Zhang

11/26/2017

Question 1: Running R scripts from the command line

a) We are interested in estimating $\hat{\beta}$, i.e the slope of the regression line.

b)

```r
library(sandwich)
```

c) The script does not give any standard errors but does provide coverage. The model gives a coverage of 0.5 while the sandwich method gives a coverage of 0.3.

d) I got the same answers as before but it was outputted in a new file.

Question 2: Running R scripts on `cox`

It look 42 seconds to run. The model gives a coverage of 0.748 and the sandwich model has a coverage of 0.9494.

Question 3: Running R scripts on the clusters

a) Running both scripts resulted in 15 text files and 15 Rdata files.

b)

n	x	model	sandwich
n20	2.003	0.7324	0.8364
n400	2.0001	0.7374	0.9474
n600	1.9999	0.7458	0.9436

Appendix

3b

```r
n = c("n20","n400","n600")

list = list()
for(i in 1:3){
  filelist = list.files(pattern = paste(n[i],".*.txt",sep =""))
  datalist = lapply(filelist, function(x) unlist(strsplit(readLines(x)[4]," ")) )
  datafr = do.call("rbind", datalist)
  frame = data.frame(n = n[i], x = as.numeric(datafr[,1]),
                   model = as.numeric(datafr[,2]),
                   sandwich = as.numeric(datafr[,3] ))
  print(frame)
  list[[i]] = frame
```

```
}

full_list = do.call("rbind", list)

full_list %>%
  group_by(n)%>%
  summarize(x = mean(x), model = mean(model), sandwich = mean(sandwich))%>%
  xtable
```