STAT 40001/MA 59800 Statistical Computing/ Computational Statistics Fall 2013
Homework 4

Name:

Due : October 3, 2013 PUID:

Instruction: Please submit your R code along with a brief write-up of the solutions. Some of the questions below can be answered with very little or no programming. However, write code that outputs the final answer and does not require any additional paper calculations.

Q.N. 1) Data from a study comparing brain size and intelligence is available on the DASL web site *http://lib.stat.cmu.edu/DASL/Datafiles/Brainsize.html.*
a) Import the Data set in R-readable format(You may first save and then import it using `read.table`)
b) Print first 5 observations.
c) It appears that there are few missing values. Identify the missing values.
d) Calculate the summary of each of the variables.

Q.N. 2) Access the data from url *http://www.stat.berkeley.edu/users/statlabs/data/babies.data* and store the information in an object named **BABIES** using the function *read.table()*
a) Create a `CLEAN` data set that removes subjects if any observations on the subject are "unknown". Note that if the values are unknown `bwt, gestation, parity, height, weight` and `smoke` are quoted as 999, 999, 9, 99, 999, and 9 respectively. Store the modified data set in an object named `CLEAN`.
b) Create side-by-side boxplots to compare the birth weights of babies for both smoking and non-smoking mothers.
c) Calculate the five number summaries of the birth weights of babies of both smoking and non-smoking mothers.

Q.N. 3) Suppose that the population of adult male bears has weights that are approximately normally distributed with average 350 lbs and standard deviation of 75 lbs. What is the probability that a randomly observed male bear weighs more than 450 lbs?

Q.N. 4) If $X \sim \chi^2_{10}$.
a) Calculate $P(X \leq 8)$
b) Calculate $P(X > 6)$
c) Calculate a so that $P(X < a) = 0.05.$

Q.N. 5) For the t-distribution we can see that as the degrees of freedom get large the density approaches the normal. To investigate , plot the standard normal density and add densities of the t-distributions with different degrees of freedom.