Homework #5

1. (5 points) Better late? An experimenter wished to determine whether there is a relationship between children's IQs and their mothers' ages when they were born. Using school records, a list was compiled of 10-year-olds whose mothers were over 45, and a second list was compiled of 10-years-olds whose mothers were 30 or under. The experimenter randomly sampled 50 children from each list and administered the Standford-Binet intelligence test to them. The IQs were found to be considerably higher for the children of older mothers, and the difference was significant beyond the 0.001 level. The experimenter concluded that a woman should postpone childbearing until later in life to ensure a high IQ for her offspring. Comment on the appropriateness of the experimenter's conclusion.

2. (15 points) Do exercise12.4.8.
Replace part (d) in textbook by:
(d) Before carrying out the study, we expected that the cholesterol levels for people with no diseases should be lower than people with disease. Conduct a hypothesis test for H_0: $\mu_1+\mu_2+\mu_3-3\mu_4=0$ by t-test. Does the conclusion changes with Bonferroni correction? Does it change with Schéffe's method? What procedure should be used here (Bonferroni, Schéffe, or no correction at all)?
Add a part (e)
(e) It is clear from the data that the group with intermittent claudication has much higher LDL cholesterol levels than other groups. Therefore we want to test H_0: $\mu_1-(\mu_2+\mu_3+\mu_4)/3=0$ by t-test. Carry out the test. Is any multiple testing adjustment procedure needed? If so, which one? If not, why not?

3. (10 points)
For the data set *airquality* contained in the R base package (there is no need to input it, see the example in Lab1), we want to know for which pairs of months the mean Ozone measurements are different. (You can use "**?airquality**" to check the description of the data set.)
(a) Do the multiple pairwise comparison test at 0.05 level, using Bonferroni, Tukey, and fdr adjustments respectively. What are your conclusions?
(b) Are the results for those three adjustments the same? If not, which you think is the more appropriate conclusion here? Why?

4. (10 points) For the data set *lowbwt* used before, systolic blood pressure measurements for 100 low birth weight infants are saved under variable sbp, gender is saved under variable sex and toxemia (during pregnancy for the child's mother) is saved under the variable tox. We wish to see whether mean systolic blood pressure is the same for low birth weight boys and girls.

(a) Use R to produce ANOVA tables: one with toxemia status as blocking variable and one without any blocking.

(b) What is the p-value for gender effect on systolic blood pressure with the blocking? What is the p-value for gender effect on systolic blood pressure without the blocking?

(c) Hypothetically, if you have a data set with randomized block design (the data set *lowbwt* is not) for testing difference among two groups. That is, each block has two subjects randomly assigned to the two groups. Is the ANOVA F-test without blocking equivalent to the independent two-sample t-test? Is the ANOVA F-test with blocking equivalent to the paired two-sample t-test?

5. (10 points) Mini-Project: measuring your response time.

Using the webpage in Lab 1 to collect 15 response times each with all four sizes of boxes: "small", "medium", "large" and "xlarge".

(a) Conduct an ANOVA to see if there is a difference between your response times for different box sizes.

(b) Conduct the pairwise comparisons between groups using HSD and LSD.