STA 570 — Spring 2012 — Dr. Charnigo

Written Assignment 4

This assignment is due on Thursday 29 March at 12 Noon. You may work in self-selected groups of two or three, submitting one copy of the assignment for the group.

[20] 1. Refer to {Diabetes.xls}, with which you have already worked in Written Assignment 2. Let p_1 denote the proportion of persons with diabetes, among those with multiple pregnancies, in the population of which the sample is representative. Let p_2 denote the proportion of persons with diabetes, among those without multiple pregnancies, in the population of which the sample is representative.

[10] a. Test $H_0 : p_1 = p_2$ against $H_1 : p_1 \neq p_2$ at level $\alpha = 0.05$ by calculating a z statistic based on the difference in sample proportions.

[10] b. Test $H_0 : p_1 = p_2$ against $H_1 : p_1 \neq p_2$ at level $\alpha = 0.05$ by calculating a χ^2 statistic based on the disparities between expected and observed values in a two-by-two contingency table.

[50] 2. Refer to {Headache.xls}, which presents data from a pilot study comparing a new pain reliever to an existing pain reliever on effectiveness in treating headaches. One hundred patients with a history of frequent headaches were directed to use the new pain reliever when they next experienced a headache and record, on a scale of 0 to 100, by what percentage they perceived the pain to be reduced one hour later. One hundred more patients, individually matched to the first one hundred patients on age and gender, were directed to use the existing pain reliever and record by what percentage they perceived the pain to be reduced one hour later. Use $\alpha = 0.05$ for all hypothesis tests.

[10] a. Create side-by-side boxplots of perceived pain reduction for the two groups.

[10] b. Apply the sign test to assess whether there is a nonzero median difference in perceived pain reduction between the two groups.

[10] c. Apply the signed rank test to assess whether there is a nonzero median difference in perceived pain reduction between the two groups.

[10] d. Proceeding as if the difference scores were normally distributed, apply the paired t-test to assess whether there is a nonzero mean difference in perceived pain reduction between the two groups.

[10] e. Do you think that the difference scores are normally distributed ? Why or why not ?

[30] 3. Refer to {FEV.xls}, which we discussed in Lecture 7. One issue with this data set is that age confounds the relationship between smoking status and FEV. We can (partially) address that issue by confining attention to participants aged 12 and older. Another issue with this data set is that FEV scores are not really normally distributed. In this exercise, we will address that issue in two ways: nonparametric testing and nonlinear transformation. Use $\alpha = 0.05$ for all hypothesis tests.

[10] a. Apply the rank sum test to assess whether median FEV differs between the two groups (smokers aged 12 and older, non-smokers aged 12 and older).

[10] b. Apply the procedure in Section 8.4 to logarithmically-transformed FEV. Do you conclude that mean logarithmically-transformed FEV differs between the two groups?

[10] c. Assuming that you could recruit one smoker for every three non-smokers, find the sample sizes from the two groups for which you would have 80% power to conclude that mean logarithmically-transformed FEV differs between the two groups.