STA 570 — Spring 2012 — Dr. Charnigo

Written Assignment 3

This assignment is due on Thursday 01 March at 12 Noon. You may work in self-selected groups of two or three, submitting one copy of the assignment for the group.

[60] 1. Refer to {Sheet1} of {MedicineLDL.xls}, with which you have already worked in Written Assignment 1. Let μ_1 and σ_1 denote the mean and standard deviation of low density lipoprotein (assumed normally distributed) for a hypothetical population of adults with a history of elevated cholesterol who are placed on active treatment. Let μ_2 and σ_2 denote the mean and standard deviation of low density lipoprotein (assumed normally distributed) for a hypothetical population of adults with a history of elevated cholesterol who are placed on placebo.

[10] a. Test $H_0 : \sigma_1^2 = 900$ against $H_1 : \sigma_1^2 \neq 900$ at level $\alpha = 0.05$.

[10] b. Test $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$ at level $\alpha = 0.05$ assuming $\sigma_1^2 = \sigma_2^2$.

[10] c. Construct a 95% confidence interval for $\mu_1 - \mu_2$ assuming $\sigma_1^2 = \sigma_2^2$. Comment on how this confidence interval relates to part b.

[10] d. Test $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$ at level $\alpha = 0.05$ assuming $\sigma_1^2 \neq \sigma_2^2$.

[10] e. Construct a 95% confidence interval for $\mu_1 - \mu_2$ assuming $\sigma_1^2 \neq \sigma_2^2$. Comment on how this confidence interval relates to part d.

[10] f. Test $H_0 : \sigma_1^2 = \sigma_2^2$ against $H_1 : \sigma_1^2 \neq \sigma_2^2$ at level $\alpha = 0.05$. Comment on the implications for parts b through e.

[10] 2. Suppose that a data analyst plans to test $H_0 : \mu_1 = \mu_2$ but waits to see whether $\bar{x} < \bar{y}$ before selecting the alternative hypothesis. If $\bar{x} < \bar{y}$, then the alternative hypothesis is chosen to be $H_1 : \mu_1 < \mu_2$. Otherwise, the alternative hypothesis is chosen to be $H_1 : \mu_1 > \mu_2$. Show that this strategy doubles the claimed significance level. For instance, if the data analyst claims the significance level to be 0.05, then the significance level is really 0.10.

[30] 3. Refer to {Data} of {Diabetes.xls}, with which you have already worked in Written Assignment 2. Let p denote the proportion of persons with multiple pregnancies, among those with diabetes, in the population of which the sample is representative. Note that p may be regarded as a conditional probability, $P($ multiple pregnancies $|$ diabetes $)$.

[10] a. Test $H_0 : p = 0.60$ against $H_1 : p \neq 0.60$ at level $\alpha = 0.05$.

[10] b. What power would you have to conduct the test in part a if the sample included 100 persons with diabetes?

[10] c. How many persons with diabetes would be needed for 90% power to conduct the test in part a?