STA 580 — Spring 2009 — Dr. Charnigo

Final Examination

Please print your name at the tops of all seven pages. You may use the backs of the pages for scratch work, but all key steps and final answers are to be clearly recorded on the fronts of the pages.

[30] 1. Refer to "Cancer" and Table 11.22 on pages 548 and 549. Let X denote the base ten logarithm of annual cigarette consumption and Y the base ten logarithm of five-year mortality. We have $L_{yy} = 0.4827$, $L_{xx} = 0.6020$, $L_{xy} = 0.5011$, $\bar{y} = -1.9438$, $\bar{x} = 0.2975$, and $n = 8$. In what follows, we consider the linear regression model $Y_i = \alpha + \beta x_i + \epsilon_i$.

[05] a. Provide the least squares estimates of α and β.

[05] b. Test $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$ at level 0.05 using an f statistic.

[05] c. Test $H_0 : \beta = 1$ against $H_1 : \beta \neq 1$ at level 0.05 using a t statistic.

[05] d. Suppose that the base ten logarithm of annual cigarette consumption is 0.60 in the time period 1970-1974. Use the linear regression model to calculate a 95% prediction interval for the base ten logarithm of five-year mortality in 1970-1974. (One may raise a question about semantics, since ordinarily we do not speak of "predicting" what happened three decades ago. Nonetheless, you can carry out the computation.)

[05] e. What fraction of variability in the base ten logarithm of five-year mortality is explained by the base ten logarithm of annual cigarette consumption?

[05] f. Calculate the Pearson correlation between X and Y. (You may calculate it directly, or you may deduce it from your answer to part e. If you use the latter approach, note that the sign of the Pearson correlation — positive or negative — must be the same as the sign of the least squares estimate for β.)

[30] 2. Refer to "Gastroenterology" and Table 12.35 on page 624. Let μ_1 through μ_3 be defined in the obvious manner. We have Total SS = 213.62 and Within SS = 193.99. Also, $f_{2,25,0.95} = 3.385$, $f_{2,25,0.975} = 4.291$, $f_{2,28,0.95} = 3.340$, $f_{2,28,0.975} = 4.221$, $f_{3,25,0.95} = 2.991$, $f_{3,25,0.975} = 3.694$, $f_{3,28,0.95} = 2.947$, and $f_{3,28,0.975} = 3.626$.

[05] a. Find Between SS.

[05] b. Find Between MS and Within MS.

[05] c. Test $H_0 : \mu_1 = \mu_2 = \mu_3$ against the complementary alternative at level 0.05. (Use a one-way analysis of variance, despite the textbook author's suggestion that the normality assumption may not hold.)

[05] d. Test $H_0 : \mu_1 = \mu_3$ and $H_0 : \mu_2 = \mu_3$ against their respective complementary alternatives at Bonferroni-adjusted level 0.05.

[05] e. Test $H_0 : 0.5\mu_1 + 0.5\mu_2 - \mu_3 = 0$ against its complementary alternative at Scheffe-adjusted level 0.05.

[05] f. Suppose that you had been given the numerical trypsin secretion scores for the 28 subjects, instead of being told only that the scores were less than 50, between 51 and 1000, or greater than 1000. What method of data analysis might you have employed then? (Just state the method; no explanation is required.)

[25] 3. Refer to "Endocrinology" and Table 13.43 on page 741. Confine your attention to the data for women with pre-existing fractures. Let p_1 denote the rate of new fractures in a hypothetical population of women with pre-existing fractures receiving raloxifene, and let p_2 denote the rate of new fractures in a hypothetical population of women with pre-existing fractures receiving placebo. (This is not the same scenario considered in the Fall 2007 Final Examination, which involved women without pre-existing fractures.)

[05] a. Report point estimates of p_1 and p_2.

[05] b. Construct a 95% confidence interval for the risk difference.

[05] c. Construct a 95% confidence interval for the relative risk.

[05] d. Construct a 95% confidence interval for the odds ratio.

[05] e. Test $H_0 : p_1 = p_2$ against $H_1 : p_1 \neq p_2$ at level 0.05 using either a z statistic or a χ^2 statistic (your choice).

[Insert joke here.]

[15] 4. Mark each of the following statements as true or false.

T F [03] a. Only if the response variable is ordinal can we employ the rank sum test to compare the median from one population to the median from another population.

T F [03] b. Only if the response variable is cardinal can we employ the rank sum test to compare the mean from one population to the mean from another population.

T F [03] c. Suppose that the survival function for a given population equals 0.70 at 1 year. From that fact alone, we may conclude that the survival function is greater than or equal to 0.70 at 6 months.

T F [03] d. Suppose that the survival function for a given population equals 0.70 at 1 year and that the survival function for another population equals 0.75 at 1 year. If performed, the log rank test will either distinguish between the two survival functions or yield a Type II error.

T F [03] e. Suppose that the Kruskal-Wallis test accepts the omnibus null hypothesis of equal medians for all populations of interest. We will follow up by using the Dunn procedure to identify pairs of population medians in which one member of the pair differs from the other.