# STA 580 — Fall 2008 — 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 "Ophthalmology" and Table 11.28 on page 553. Let $X$ denote the 1999 serum-lutein value, and let $Y$ denote the 2003 serum-lutein value. We have $L_{yy} = 64.50$, $L_{xx} = 20.89$, $L_{xy} = 34.97$, $\bar{y} = 6.878$, $\bar{x} = 3.511$, and $n = 9$. 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 $\alpha$ and $\beta$.

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

[05] c. Test $H_0 : \beta = 2$ against $H_1 : \beta \neq 2$ at level 0.05 using a $t$ statistic.

[05] d. Provide a 95% prediction interval for the 2003 serum-lutein value of another patient (not in Table 11.28) whose 1999 serum-lutein value was 5.7.

[05] e. What fraction of variability in 2003 serum-lutein values is accounted for by 1999 serum-lutein values?

[05] f. Even though we are viewing $X$ as a predictor of $Y$, we can still go through the mechanics of calculating a Pearson correlation. Please do this and briefly comment on how the Pearson correlation relates to your answer to part e.

[30] 2. Refer to "Cardiovascular Disease" and Table 12.39 on pages 627 and 628. Let $\mu_1$ through $\mu_4$ be defined in the obvious manner. We have Between SS = 344.1 and Within SS = 2212.9. Also, $t_{142,0.975} = 1.977$, $t_{142,0.9875} = 2.265$, $t_{142,0.9917} = 2.424$, $t_{146,0.975} = 1.976$, $t_{146,0.9875} = 2.264$, $t_{146,0.9917} = 2.423$, $f_{3,142,0.95} = 2.668$, $f_{4,142,0.95} = 2.435$, $f_{3,146,0.95} = 2.667$, and $f_{4,146,0.95} = 2.434$.

[05] a. Find Total SS.

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

[05] c. Test $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ against the complementary alternative at level 0.05.

[05] d. Test $H_0 : \mu_1 = \mu_2$ and $H_0 : \mu_3 = \mu_4$ against their respective complementary alternatives at Bonferroni-adjusted level 0.05.

[05] e. Test $H_0 : \mu_1 - \mu_2 - \mu_3 + \mu_4 = 0$ against its complementary alternative at Scheffe-adjusted level 0.05.

[05] f. One could also perform a Scheffe adjustment in part d instead of a Bonferroni adjustment. Unlike the Bonferroni adjustment, the Scheffe adjustment would remain valid if I were to come back tomorrow and ask you to do one more test in part d. Hence, with a Scheffe adjustment you would not have to 'redo' the original tests if I were to come back tomorrow and ask you to do one more test, whereas with a Bonferroni adjustment you would have to 'redo' the original tests. Given this advantage of a Scheffe adjustment, briefly explain why a data analyst may still prefer to use a Bonferroni adjustment.

[25] 3. Refer to "Hypertension" on page 735. Let $p_1$ denote the rate of elevated systolic blood pressure among people with high blood lead levels, and let $p_2$ denote the rate of elevated systolic blood pressure among people with low blood lead levels. Taking into account both men and women, we have the following contingency table.

| Sample | Elevated SBP | Non-Elevated SBP | Row Total |
|---|---|---|---|
| High Blood Lead | 17 | 585 | 602 |
| Low Blood Lead | 10 | 1108 | 1118 |
| Column Total | 27 | 1693 | 1720 |

[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 $\chi^2$ statistic (your choice).

A biostatistician walks into $\bar{a}$.

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

T F [03] a. If we have an ordinal response variable and wish to compare the mean from one population to the mean from another population, we may draw independent samples from the two populations and then employ the signed rank test.

T F [03] b. If we have a cardinal response variable and wish to compare the mean from one population to the mean from another population, we may draw independent samples from the two populations and then employ the rank sum test.

T F [03] c. If the survival function for a given population equals 0.80 at 1 year, then 20% of the people in that population experience the event of interest within 1 year.

T F [03] d. If 20% of the people in a given population experience the event of interest within 1 year, then a Kaplan-Meier estimate of the survival function constructed from a sample from that population will equal 0.80 at 1 year.

T F [03] e. The Dunn procedure may be used to perform follow-up tests concerning pairs of population medians after the Kruskal-Wallis test rejects the omnibus null hypothesis of equal medians for all populations of interest.