STATISTICS 210 Midterm Solutions – Fall 2008

1. **TV and violence / ANOVA**

   (a) The population parameters in this case are the mean violent behavior scores for the "populations" (or maybe subpopulations would be better) of children that watch a given amount of TV per day. Note the population refers to the children in the groups but the parameter is the numerical summary of the population, in this case mean violent behavior score.

   (b) First, fill in the table by noting there are 5 groups, so d.f. for groups is 5-1=4 and the d.f. for error is 2267-5=2262. The SS(Error) $= \sum (n_i - 1)s_i^2 = 14586.6$ and the SS(Group) $= \sum n_i * (\bar{Y}_i - \bar{Y}_{..})^2 = 601.14$. You should only have had to calculate one because I gave you the total. Note the overall mean (used in computing SS(Groups) can be computed in two different ways (weighted and unweighted) because under $H_o$ they yield the same expected value. I prefer the weighted calculation but either was OK here. Now we want to test $H_o : \mu_{<1} = \mu_{1-2} = \mu_{3-4} = \mu_{5-6} = \mu_{>6}$ vs the alternative $H_a$ : all means are not equal. $F = MS(Group)/MS(Error) = (601.14/4)/(14586.6/2262) = 23.3$. The $P$-value from the table is $< .001$ (in fact, we can learn from a computer that it is much smaller!). This means the observed data are extremely unlikely under $H_o$ and thus we would reject $H_o$ and conclude there are group differences.

   (c) The intent of this question was for you to think about the distribution of scores **within** each group rather than to compare s.d.'s across groups. The s.d. is large compared to the mean (and the score is nonnegative) which means the distribution of scores must be skewed and hence not normal. Fortunately the sample sizes are quite large so this is probably not a problem for our analysis. It is true that the s.d.'s are nearly equal so that assumption of constant variance is probably OK too, but you did not need to address the s.d.'s for full credit.

   (d) Contrast

      i. The contrast of interest is $\gamma = -2\mu_{<1} - 1\mu_{1-2} + 0\mu_{3-4} + 1\mu_{5-6} + 2\mu_{>6}$. The estimated contrast is $\hat{\gamma} = 2.69$ and it's estimated standard error is $\sqrt{MSE * \sum_i c_i^2/n_i} = \sqrt{6.45 * .0278} = .423$. Then we can test $H_o : \gamma = 0$ vs $H_a : \gamma > 0$ using $t = \hat{\gamma}/s.e.(\hat{\gamma}) = 2.69/.423 = 6.36$. The $P$-value (1-sided) from t table with 2262 d.f. is $P < .0005$ (in fact, we can learn from computer that it's much lower) so we reject $H_o$ and decide the data do support the linear contrast. Thus, there is a linear pattern in the means (note, it is not perfectly linear but there is a linear component). A couple of comments here. First, a confidence interval is not a test. When you are asked for a test you should give one. Second, this seemed like a natural place for a one-sided alternative. You got full credit for either one-sided or two-sided alternatives as long as you were consistent and defined your p-value in the same way as $H_a$. Lastly, note that you can't make a hypothesis about $\hat{\gamma}$ – we know this quantity!

      ii. My intention here was for you to see what happens if the sample mean for the middle group changed without changing anything else. It turns out that neither the sample contrast or the s.e. would change so we would get exactly the same test result! This has always bothered me a little. It is correct though. One way to understand this is to note that if you change that middle mean the data DO STILL SUPPORT the linear pattern BUT you will also find a significant (perhaps more significant) quadratic component. Several people argued that the s.d. would probably change if the mean did – this is a good point but it's not necessarily true here where the scores are limited to 0 - 15.

2. **Sleep and study / regression** - There was a typo on the exam. The regression variable was listed as "sleep" when it was supposed to be "study". Almost everyone (more than 90%) interpreted it the way I intended. I gave credit for either but the answer here reflects my intention that "study" be the explanatory variable.

   (a) The average amount of sleep decreases about 1/4 hour for every 1 hour increase in study time. A key point is that the interpretation of the regression coefficient is about the population average or expected amount of sleep rather than about an individual's amount of sleep.

   (b) You are given $\hat{\beta}_1 = -.2555$ and $s.e.(\hat{\beta}_1) = .0879$. All you need is $t_{62,.975} \approx t_{60,.976} = 2.000$ to find that the 95% CI is $-.2555 \pm 2 * .0879 = (-.43, -.08)$.

   (c) Because 0 is not in the CI we know that a two-sided test of the hypothesis that the slope is zero would reject at the .05 level.

   (d) The prediction for an individual with 3 hours of study time is $7.8659 - 3 * .2555 = 7.0994 \approx 7.10$ hours. The approximation that I was looking for was the recognition that the prediction standard error would include the "extra" $s_e$ term and thus would be approximately 1.6363. This would give an approximate PI of $7.10 \pm 2 * 1.64 = (3.82, 10.38)$. This ends up being fairly precise if you check the formula because 3 study hours is close to the sample mean of 2.99 – the prediction standard error is $\approx 1.6363 * \sqrt{65/64} = 1.65$.

3. **Intelligence and smoking / two-sample**

(a) Let $\mu_{smok}$ and $\mu_{non}$ denote the mean 4-year old IQ scores for the population of smoking and nonsmoking mothers. The two s.d.'s are close so we can use the pooled procedure. Then the 95% CI for $\mu_{non} - \mu_{smok}$ is $\bar{Y}_{non} - \bar{Y}_{smok} \pm t_{111,.975} s_p \sqrt{1/n_{non} + 1/n_{smok}}$. The pooled s.d. is $\sqrt{(46 * 12.9^2 + 65 * 14.0^2)/111} = 13.55$ and the CI is $10.2 \pm 2.000 * 13.55 * 0.19 = (5.0, 15.4)$. Note that it is best to round $t$ d.f. down if you don't have a calculator/computer nearby so that here you would use 60 d.f.. You were asked for a one sentence summary: We are 95% confident that the difference in mean 4-year old IQ scores is between 5 and 15.4 points higher for the population of children born of non-smoking mothers than for the population of children born of smoking mothers. The correct interpretation of a CI is NOT to report on a significance test! If I wanted a test, then I would have asked for one.

(b) The study does not prove a **causal** link because it is an observational study (and not a randomized study). There are many possible confounding factors (e.g., education, socioeconomic status, alcohol use) that may differ between the two groups and may explain the observed difference in mean IQ score. A randomized study would tend to balance these other factors. You needed to at least mention confounding for full credit.

(c) Most people identified the small sample size as a limiting factor in this attempted replication. Several also mentioned possible differences in the population. Good stuff. I was hoping (if time had been less of an issue) that people would try to compare effect sizes. In the large study you have the data from the 5-year olds so can calculate $eff_1 = (113.8 - 106.9)/13.5 = 0.51$. For the smaller study, recall from the HW that $t = eff * \sqrt{n_1 * n_2/(n_1 + n_2)} = eff * \sqrt{6}$ so that $t = 1.25$ yields $eff_2 = 1.25/\sqrt{6} = 0.51$. Thus in fact the second study provides an exact replication of the effect size in the first study. It is purely sample size that keeps it from being significant.