Homework #7: Teasing Relationships from Univariate Data

- **Guidelines:** Hand in your output file. (For this one, do, regardless of what I've said otherwise. You'll be doing some advanced stuff, and I need to see that you did it correctly.) Write answers in sentence form, and be explicit and complete in explaining your answers. Show all work done, including any calculations, and explain all of the steps involved. Be certain to report all statistics in terms of some unit of analysis – for example, years old, or dollars earned – and be clear to distinguish between original units and standard errors.

- **Work Alone:** You may *plan* your work in study groups. You may even work together to perform the entire analysis on a different data set. But *once you begin analyzing data, you are on your own and may only consult your book, your notes, and the website!* You may of course ask me anything you wish, although some questions I can't answer. Email is quickest and easiest: egodard@csun.edu

- **Study Topic:** Based on your recently acquired statistical skills, you've been hired as a consultant by a local agency. You have been asked to assess whether intergenerational transmission of educational attainment is the same for men and women – that is, whether women follow in their mothers' footsteps as much as men follow in their fathers' footsteps. You will need to conduct the following analyses, using data from the 1984 General Social Survey (GSS84.SAV):

- **Find the data:** Identify variables that measure the education level of the respondents, their fathers, and their mothers. You do *not* need to *discuss* univariate analysis of these variables for this assignment. But you should always look at frequency distributions (and histograms, if appropriate) before conducting analysis.

- **Clean the data:** Create two new variables, each of which measures the respondents' improvement over one of their parents in terms of education – i.e. how much *more* education they received than their mother and father. (Hint: To compute how different you and I are in age, you might use AGEDIFF = MYAGE – YOURAGE. Don't use age. This assignment isn't about age. That was a hint about computing, not a demonstration.)

1) **Compute confidence intervals:** Using SPSS output, construct confidence intervals for the mean improvement for each of four groups. (The first two are just your two new variables, for the full sample; the others are each variable, for one subsample. You could get the 3rd and 4th using SELECT IF twice, or at the same time using the descriptives table from ANOVA or t-test output.) Be sure to *interpret the results!*
    - all respondents compared to their fathers (using one of the new variables)
    - all respondents compared to their mothers (using the other of the new variables)
    - male respondents compared to their fathers (using one of the new ones, just for men)
    - female respondents compared to their mothers (using the other one, just for women)

2) **Perform four tests:** (Complete all five steps for each test, include carefully specified hypotheses!)
    a) An hypothesis test for whether the sample is balanced in terms of gender, or if there are significantly more than 50% men or women. (This isn't about any education variables, though case totals there could help.)
    b) A difference of proportions test for the proportion of men who improved over their fathers versus the proportion of women who improved over their mothers. (Hint: Use the cumulative frequencies column.)
    c) An hypothesis test for whether women improve over their mothers more than the sample as a whole does
    d) A difference of means test for whether men improve over their fathers as much as women improve over their mothers

3) **Make conclusions**: State (in a few sentences, at least) a conclusion about sex differences in intergenerational change in education, taking account of any relevant factors (sample sizes, representativeness of sample proportions, confidence levels, satisfaction of assumptions made, etc.) that you have encountered during your study.

Optional Extra Credit Homework

Want more practice? Do problems 8, 16, 22, and 24 from chapter 7 (pp. 206-214 in the Agresti/Finlay text, 5th edition) for extra credit. I'll grade it as an additional assignment, so you'll have 7 instead of 6. (That is, you'll have 10 and I'll drop the lowest three leaving 7, rather than 9 for which dropping the lowest three leaves 6. This is an *extra* homework, *not* a replacement for something else you skip or do poorly on.) If your homework grades are skewed, you could minimize the effect of an outlier on the low end. In no case will attempting this extra credit opportunity *hurt* your grade: Your grade on it will only be counted if it will raise your homework grade, and I'll calculate it both ways, to confirm.