Homework 3
Due Thur, Mar 11, 10 AM (pdf format please)

For all questions, please show your work or include a copy of the output, whichever is relevant. Please write your answers in report form, as if you were describing results in a published study. Most responses should not be longer than one paragraph. The data for each problem can be downloaded from the website: [http://web.pdx.edu/~newsomj/data.htm](http://web.pdx.edu/~newsomj/data.htm). All answers should be in your own words.

1. Use the school survey data (npc.sav) from HW 2 for this problem. Using SPSS and R, conduct a multiple logistic regression analysis predicting gun use with race, neighborhood support, neighborhood erosion, and gang membership. Report and interpret your results, including the overall fit, the psuedo-$R^2$, all of the odds ratios, significance tests, and confidence intervals.

2. For Problems 2a-c below, use your own data or the drug court data described below. If you use your own data set, please provide a paragraph description of the study and variables you will use for the analysis. If you do not have access to suitable data, download the drug court data set (drug.sav) from the website at [http://web.pdx.edu/~newsomj/data.htm](http://web.pdx.edu/~newsomj/data.htm). The drug court data set comes from a survey of individuals arrested for drunk driving in New Mexico and contains a subset of the variables reported by the defendant, including age (age), gender (gender, 0=male, 1=female), whether currently employed (work, 0=not employed, 1=employed), whether defendants agree that the term adventurous describes them (advent, 1="strongly disagree" to 5="strongly agree"), whether the defendants agree they have family support (support, 1="strongly disagree" to 5="strongly agree"), whether the defendants believe they are guilty of drinking while intoxicated (dwi, 0="not guilty", 1="somewhat guilty", 2="very guilty"), and the number of prior convictions (priors).

   a. If you are using your own data, chose a dependent variable that has 3 or 4 ordinal values and test an ordinal logistic model in SPSS or R using four predictors. If using the drug court data, use age, gender, employment, and adventurousness as predictors of the defendants perceived guilty (dwi) in an ordinal logistic model. Report and interpret your results. Be sure to report the likelihood ratio chi-square, pseudo-$R^2$, the regression coefficients, odds ratios, coefficient significance, and confidence intervals.

   b. Test the same model as the model above in SPSS or R but use an ordinal probit analysis. Report and interpret your results. Be sure to report the likelihood ratio chi-square, pseudo-$R^2$, unstandardized and standardized regression coefficients, coefficient significance, and confidence intervals.

   c. For this question, use your own data if you have a count variable or use the drug court data. If you are using your own data, chose a dependent variable that is a count and test a negative binomial regression model in SPSS or R using two or more predictors. If using the drug court data, use gender, work, and family support to predict the number of priors in a negative binomial model. Be sure to report the likelihood ratio chi-square, the regression coefficients, coefficient significance, and confidence intervals. Compare the two $-2 \times \log$ likelihood value- to one obtained from a Poission regression with the same variables to see if there is was an overdispersion problem.

3. For this problem, use your own data or use the structured play data set from HW1 (play.sav). If you use your own data, you should compare at least two groups on at least three dependent variables (which should be at least modestly positively correlated). If you use the structured play data, first rescore the bullying variable by subtracting each score from 3 (e.g., compute bullyingr=3 – bullying), then use SPSS or R to conduct a MANOVA to compare the two treatment groups (treat) on the dependent variables bullying (reversed), playsup, and emoreg. Report the
multivariate results, including the F-value, significance, and the univariate tests. Be sure to include the univariate means of each and indicate the direction of the differences to describe your results.

4. Read one of the following articles (password protected copies are available from the class website http://web.pdx.edu/~newsomj/) and write two paragraphs summarizing the article. First, describe the study design (e.g., randomized experiment, non-equivalent control group design, cross-sectional survey; for a quick refresher, see http://sphweb.bumc.bu.edu/otlt/mph-modules/programevaluation/ProgramEvaluation7.html) and purpose of the study in your own words. Be sure to include who/what was studied (e.g., who were the participants?) and the number of cases. Then, choose one statistical test used in the article that you have learned about in the course in this section (i.e., logistic regression, ordinal regression, count regression, multivariate analyses), and, in your own words, describe the hypothesis that is being tested, the results obtained, and what the findings mean. Be sure to include the relevant statistical values and whether the results were significant. Write your paragraphs as if you were describing results in a published article and reporting someone else’s results as in a review article.


van Beek, I., Taris, T. W., & Schaufeli, W. B. (2011). Workaholic and work engaged employees: Dead ringers or worlds apart?. Journal of Occupational Health Psychology, 16(4), 468.