ICPSRCDA18 Assignment 6: the multinomial logit model

Your Name: Points received: ____ out of 115

Your dependent variable **must** have value labels where the first letter or number of the label differs for each outcome category. To help me match your results to the output, **you must** include in parentheses the first letter from the value label for each outcome when you refer to it (e.g., The odds of reporting poor health {p} compared to excellent health {e} are 1.4 times greater for men than women, holding all other variables constant). Additionally, to avoid potential problems, your outcome variable Y should be coded 1, 2, 3, …

1. ___ of 5: Choose one of the data sets available for class. Pick or construct a nominal dependent variable and at least **four** independent variables: a binary variable B; a continuous variable C; and **two additional** variables X and W. (Note: You can use the same model from your OLM assignment **IF** the parallel regression assumption was violated). Keep only the variables you will be using and drop all missing cases (listwise deletion). (Note: C and B must be statistically significant—see #5 below; if not, you must find new variables.)

2. ___ of 5: Demonstrate that the data are clean by including the output from the following commands (or their R equivalents):

   `. codebook, compact`

   `. sum`

3. ___ of 5. Present a simple table describing your variables. Include the distribution of the dependent variable and the value labels and letters you will use in your graphs. This should look similar to the table on slide 20 of the Lecture Notes.

4. ___ of 5: Estimate the MNLM of Y on C, B, X, and W. Include the full output that includes all factor change coefficients for all comparisons (e.g. from `listcoef, help`), not just the minimal set of coefficients.

5. ___ of 5: Compute both a Wald and LR test that the effect of B is zero; and then do the same for C. Write up the results of the Wald test **OR** the LR test (not both) of C as though it were part of a published paper.

6. ___ of 10: Create an **odds ratio plot** for variables C and B ONLY. Include lines to indicate statistical significance at the 0.05 level. Use the note option to provide a key to the meaning of the symbols. Show the plot.

7. ___ of 10: **In outline form**, describe the pattern of effects of C and B in general, substantive terms. Use the graph from question 6 to help you uncover the overall pattern, but do NOT refer to the graph (i.e., describe the pattern to someone who will not see the graph). Do not discuss specific numbers, only the pattern of effects, the relative magnitudes, and significance.

8. ___ of 10: Compute the discrete change in probabilities for C and B with other variables held at some value you find interesting or useful. Only show the output for variables C and B. **Plot the discrete changes** for C and B that you find interesting or useful. Use the note option to provide a key to the symbols in the plot. Show the plot.

9. ___ of 10: Compute the discrete change in probabilities for C and B *with the other variables held at location different from that* in question 8. Only show the output for variables C and B. **Plot the discrete changes** for C and B that you find interesting or useful. Use the note option to provide a key to the symbols in the plot. Show the plot.

10. ___ of 10: **In outline form**, describe the pattern of discrete change effects of C and B in general, substantive terms. Use the graphs from questions 8 and 9 to help you uncover the overall pattern, but do NOT refer to the graph (i.e., describe the pattern to someone who will not see the graph). Do not discuss specific numbers, only the pattern of effects, the relative magnitudes, and significance.

11. ___ of 20: Using the outlines from questions 7 and 10, write a paragraph telling the story of your results for C and B as it might appear in a published paper. Assume that you are NOT including the plots in your paper, but instead are using the plots to uncover the patterns. This paragraph should expand on your outline by including magnitudes of specific effects as well as information on statistical significance, drawn from the output given by `mlogtest, listcoef, mchange,` and `margins + mlincom,` or their R equivalents. This paragraph should be concise but complete.

12. ___ of 10: What do you learn substantively from the information on discrete changes that was not clear from the information on odds ratios? What do you learn substantively from the information on odds ratios that was not clear from the information on discrete changes?

13. ___ of 10: My assessment of the overall effectiveness of your answers.