ICPSRCDA17 Assignment 5: the ordinal logit model

Your Name: Points received: ____ out of 125

In this assignment, you will estimate the ordered logit model (OLM) and interpret the results.

1. ___ of 5: Using one of the data sets for class, choose or create a substantively reasonable, ordinal dependent variable with at least four categories (I'll refer to it as Y). Choose or create at least four independent variables: a binary variable B; a continuous variable C; and two additional variables X and W. Keep only the variables you will be using and drop all missing cases (listwise deletion). Demonstrate that the data are clean by including output from `codebook, compact` and `sum`.

2. ___ of 5: Present a simple table describing your variables. Include the distribution of the dependent variable. This table should look similar to the table on slide 41 of the Lecture Notes.

3. ___ of 5: Estimate the OLM regressing Y on C, B, X and W. Include the output from `ologit`.

4. ___ of 5: Use `listcoef` to compute the factor (or percent change) in the odds. Highlight the numbers used to answer questions 5 and 6 and indicate which question they are used for.

5. ___ of 5: Using the factor change (or percent change) coefficient, describe the effect of C as though it is part of a published paper. Use either standardized or unstandardized coefficients, and include information on statistical significance.

6. ___ of 5: Using the factor change (or percent change) coefficient, describe the effect of B as though it is part of a published paper. Include information on statistical significance.

7. ___ of 5: Using `mchange`, compute the discrete change coefficients for C and B (but only these two variables!) with covariates held at their mean. Include this output.

8. ___ of 5: Highlight the appropriate discrete change coefficients for B from the output in 7 above and reproduce these values using `margins` and `mlincom`. Include this output.

9. ___ of 10: Describe the discrete change for B as though it were part of a published paper. It is not necessary to interpret the discrete change for all categories of the dependent variable (although you should interpret more than just one). Instead of listing the effect of B on each outcome category, paint an overall picture of the effect. Include information on statistical significance in your interpretation.

10. ___ of 5: Choose **one** appropriate form of the discrete change coefficients (e.g., change from min to max, unit change, standard deviation change) for C from the output in 7 above, highlight these numbers, and reproduce these values using `margins` and `mlincom`. Include this output.

11. ___ of 10: Describe the discrete change for C as though it were part of a published paper. It is not necessary to interpret the discrete change for all categories of the dependent variable (although you should interpret more than one). Instead of listing the effect of C on each outcome category, paint an overall picture of the effect. Include information on statistical significance in your interpretation.

12. ___ of 10: Use `mgen` to calculate the predicted probability of Y across the range of C. Hold the other variables at their means. Use the `graph` command to construct two plots, one of predicted probabilities and one of predicted cumulative probabilities (similar to those on slide 70 of the Lecture Notes). Be certain that the reader can easily determine which region of the graph corresponds to which categories. Include both the plots and the commands used to create the plots.

13. ___ of 15: Using **one** of the plots from question 12, describe the effect of C on Pr(Y). This should read as though it were part of a published paper and should include the magnitude of effects (obtained using `mchange` or `margins`) and information on statistical significance where necessary (obtained using `margins` and `mlincom`). It is not necessary to list or interpret the effect of C on each category of the dependent variable. Rather, paint an overall picture of the effect.

14. ___ of 10: Use `brant` to test the parallel regression assumption. Include the output and highlight the numbers used. Describe the results of the test as though it were part of a published paper. Report on both the omnibus test & specific variables.

15. ___ of 15: In 2002 the International Social Survey Program fielded the "Family and Changing Gender Roles III" module that included the following question: "A working mother can establish just as warm and secure a relationship with her children as a mother who does not work." The response categories are: 1. Strongly agree; 2. Agree; 3. Neither agree nor disagree; 4. Disagree; 5. Strongly disagree. However, the codebook reports that in the US the category '4. Disagree' is omitted. Give simple and direct answers to the following:

a) Assume that the question was asked with all categories presented, but that a programming error led to categories 4 and 5 being combined. Would it be reasonable to compare the factor change coefficients from a model estimated in the US to those from other countries where the categories were not collapsed? What about the discrete change coefficients?

b) Now assume that the error was made by forgetting to include 'disagree' on the survey itself (i.e., respondents did not get a chance to choose disagree.). Would it be reasonable to compare the factor change coefficients from a model estimated in the US to those from other countries? What about the discrete change coefficients?

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