Assignment 07

Analysis of Covariance Using Multiple Regression

This assignment is intended to give you more experience fitting and interpreting regression models with categorical predictors. Submit your responses to each of the questions below in a printed document. All graphics should be resized so that they do not take up more room than necessary and also should have an appropriate caption. This assignment is worth 15 points. (Each question is worth 1 point unless otherwise noted.)

For this assignment, you will again examine whether there are differences in course evaluation scores—and thus on earnings differences—between native and non-native English speaking professors. To do so, you will use the file *evaluations.csv* (see the data codebook).

Unadjusted Group Differences Model: ANOVA

Fit a regression model using the dummy-coded `native_english` predictor to explain variation in course evaluation scores. Use the `glance()` and `tidy()` functions to examine the output.

1. Write the fitted regression equation.

2. Interpret the intercept coefficient.

3. Interpret the slope coefficient.

4. In terms of means (not betas), write the null hypothesis (using mathematical notation) associated with the *t*-test of the slope? Be specific.

5. Based on results of the *t*-test for the slope, what do you conclude about differences in evaluation scores between native and non-native English speaking professors?

6. Use the fitted regression equation to estimate (a) the mean course rating for native and (b) the mean course rating for non-native English speakers. Show your work.

Adjusted Group Differences Model: ANCOVA

Now, suppose you want to examine differences in course evaluation scores between native and non-native English speakers, but this time you want to control for differences in professors' beauty ratings and the number of courses for which the professor has evaluations. Fit this model and use the `glance()` and `tidy()` functions to examine the output.

7. Write the fitted regression equation.

8. Interpret the fitted regression coefficient for `native_english`.

9. Compare the size and direction of the difference in course evaluation scores between native and non-native English speakers in the adjusted model to those from the unadjusted model. How do they compare? Also compare the uncertainty in the estimates.

10. Write the fitted regression equation for native English speakers. (Note: This equation should only include the effects of beauty and the number of courses for which the professor has evaluations.)

11. Write the fitted regression equation for non-native English speakers. (Note: This equation should only include the effects of beauty and the number of courses for which the professor has evaluations.)

12. Compute the *adjusted* mean course rating for native and non-native English speakers (based on professors having an average beauty and an average number of courses). Show your work.

13. Create the density plot of the marginal distribution of the standardized residuals from the ANCOVA model. Add the confidence envelope for the normal distribution. Explain whether or not this plot suggests problems about meeting the normality assumption.

14. Create the scatterplot of the standardized residuals versus the fitted values from the ANCOVA model. In the plot identify observation with extreme residuals (≤ -3 or ≥ 3) by indicating the row number of that observation in the plot.

15. Explain whether or not this plot suggests problems about meeting the linearity and homogeneity of variance assumptions.