Assignment 06

Analysis of Covariance Using Multiple Regression

Please submit your responses to each of the questions below in a printed document. Also, please adhere to the following guidelines for further formatting your assignment:

- All graphics should be resized so that they do not take up more room than necessary and should have an appropriate caption and labels.
- Any typed mathematics (equations, matrices, vectors, etc.) should be appropriately typeset within the document using Equation Editor, Markdown, or LaTeX.

This assignment is worth 15 points. Each question is worth one point unless otherwise noted.

For this assignment, you will use the file *beauty.csv*. This file contains data collected from student evaluations of instructors' beauty and teaching quality for several courses at the University of Texas. The teaching evaluations were conducted at the end of the semester, and the beauty judgments were made later, by six students who had not attended the classes and were not aware of the course evaluations. The variables are:

- `prof`: Professor ID number
- `avgeval`: Average course rating
- `btystdave`: Measure of the professor's beauty composed of the average score on six standardized beauty ratings
- `tenured`: 0 = non-tenured; 1 = tenured
- `nonenglish`: 0 = native English speaker; 1 = non-native English speaker
- `age`: Professor's age (in years)
- `female`: 0 = male; 1 = female
- `students`: Number of students enrolled in the course
- `percentevaluating`: Percentage of enrolled students who completed an evaluation

These source of these data is: Hamermesh, D. S. & Parker, A. M. (2005). Beauty in the classroom: Instructors' pulchritude and putative pedagogical productivity. *Economics of Education Review, 24*, 369–376. The data were made available by: Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.

Unadjusted Group Differences Model: ANOVA

Use the `lm()` function to fit a regression model with `avgeval` as the outcome and `nonenglish` as the predictor. Use the `summary()` function 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), what is the null hypothesis associated with the t-test of the slope? Be specific.

5. Based on the t-test of the slope, what do you conclude for the two-group comparison?

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. **(2pts.)**

Adjusted Group Differences Model: ANCOVA

Now, suppose you want to examine differences between native and non-native English speakers, but this time you want to control for differences in beauty (`btystdave`) and the percentage of students who completed a course evaluation (`percentevaluating`). Fit this model using `lm()` and use the `summary()` function to examine the output.

7. Write the fitted regression equation.

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

9. Compare the `nonenglish` coefficient and standard error in the adjusted model to those from the unadjusted model. How do they compare?

10. Write the fitted regression equation for native English speakers. (Note: This equation should only include the predictors `btystdave` and `percentevaluating`.)

11. Write the fitted regression equation for non-native English speakers. (Note: This equation should only include the predictors `btystdave` and `percentevaluating`.)

12. Compute the *adjusted* mean course rating for native and non-native English speakers (based on professors having an average beauty and an average percentage of students who complete course evalutions). Show your work. **(2pts.)**

13. Using the `ggplot()` function, create a plot of the fitted model showing the relationship between the percentage of students evaluating the course and average course rating, controlling for the beauty of the instructor. Be sure this plot includes the fitted lines for both native and non-native English speakers. Be sure to differentiate the lines, using different line styles (not color—unless you plan to print in color). Include the plot in your word-processed document.