Psychology 202a

Advanced Psychological Statistics

Sixth homework assignment, 11/10/2020 (due 11/17/2020).

In this assignment, we will work with the same dataset you sampled and used in Homework 1. If you're unsure where you stored your specific dataset, download yours off of CatCourses through the same assignment where you uploaded it. Here is a link to the codebook for the Statlab data set.

In past homework, we have used scores of verbal intelligence from the Peabody and scores of non-verbal intelligence from the Raven. Our dataset also contains the variable for each mother's age at her child's birth. There are studies (for instance Cohen, 2014) that show children of older mothers are more likely to develop cognitive disabilities (autism, for example), but others that suggest the greater maturity of older mothers helps with development (Sutcliffe, 2012). Hence, the expectation of relation between intelligence and maternal age is an interesting question that's worth examining. Specifically, we would like to examine the relationship between maternal age and Raven scores.

Run the multiple regression of Raven on Peabody and maternal age. (That is, use Peabody and mother's age at her child's birth as two predictors in a single multiple regression equation, and use Raven as the dependent measure.)

Check the regression assumptions. You should consider doing a normal Q-Q plot of the residuals. You will certainly want to examine a plot of residuals against predicted values when you assess homoscedasticity. In addition, check raw plots of Raven against each of the predictors. Write a brief commentary that summarizes your conclusions about the regression assumptions.

Comment on the regression results. Is the model as a whole significant? Are any of the individual predictors significant? Would it be appropriate for you to continue to include maternal age as a part of your multiple regression equation? Why or why not? Ignoring questions of significance, what proportion of variability in Raven is "explained" by the model? What is the predicted mean of the hypothetical population of children who scored 81 on the Peabody and had mothers age 25 at her child's birth?

Now, produce an added variable plot that shows the partial relationship between Raven and Peabody, controlling for the maternal age. Compare the plot with the raw plot of Raven against Peabody. What differences do you observe? Can you explain them?

Regress the Raven residuals that you used in the added variable plot on the Peabody residuals, and confirm that the slope is the same as the slope you observed in the multiple regression. Is the *t*-test for the null hypothesis that the slope is zero identical to the corresponding test in the multiple regression? Can you explain why, or why not?