PA 9950 Quantitative Methods

Problem Set #2

This homework is to accompany program #1 which you should all receive via email.

Answer all questions fully, please be sure to turn in a log file with your assignment. You should write up your results as though you are providing a reply to a reviewer.

Questions for Section 1:

1. Be sure to calculate the means before you estimate the regression. Re-scale the cost variable to aid in interpretation.
2. Examine and interpret the coefficient estimates of the log(salary) regression. Discuss the difference between the estimation sample and the full sample. To what should we ascribe the difference?
3. Discuss the hypothesis that all the parameter estimates are jointly equal to zero.
4. Explain the variance inflation factor. How is it calculated? When is it cause for concern?
5. Estimate two separate models dropping faculty in one and then class size in the other. Are the variables statistically significant?
6. Discuss the difference between the hypotheses:
   1. test clsize=faculty=0
   2. test clsize-faculty=0
   3. test clsize+faculty=0
7. Manipulate the variables in the regression to provide a direct test of these hypotheses without using the test command. Compare your results.
8. Predict both the residual and yhat for this model
   1. Describe how the kdensity command works and visually interpret the results of the kernel density plots.
   2. Test whether they are normally distributed using sktest. What is H0?
9. Define skewness and kurtosis. Discuss how this test works. Discuss the implications for inference based on this test!
10. Predict the median starting salary in both logs and levels.
    1. Discuss how you would take into consideration E[exp()] and the fact that the error may not be normally distributed.