Assignment 04

Probability Distributions

This goal of this assignment is to give you experience using probability distributions in a regression analysis. 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. If you are using Markdown, all syntax should be hidden (i.e., not displayed) unless specifically asked for. Any messages or warnings produced from loading packages should also be hidden. This assignment is worth 10 points. (Each question is worth 1 point unless otherwise noted.)

In this assignment, you will use the data from the file *evaluations.csv* (see the data codebook to explain variation in students' course ratings.

Part I: Model-Level Inference

Fit a regression model that uses professor's beauty, professor's age, and professor's sex to explain variation in course ratings. Do not trim this model; leave all predictors in the model whether they are statistically significant or not.

1. Partition the total sum of squared error and degrees of freedom into model and error components.
2. Use the partitioning from Question #1 to compute the mean squares for the model and error. Show your work.
3. Compute the observed F-statistic. Show your work.
4. Use the cumulative density functionality in R to find the p-value to test whether $\rho^2 = 0$ using the observed F-value you computed in Question #3.
5. Create an ANOVA table that gives the model and error partitioning for the sum of squares and degrees of freedom. Also include the mean squares, the F-statistic and the p-value. The formatting of this table will look similar to the output from the `anova()` function, except the partitioning will only be between model and error.

Part II: Coefficient-Level Inference

For all of these questions, use the partial effect of age from the fitted model.

6. Compute the observed t-statistic associated with the partial effect of age. Show your work.
7. Use the cumulative density function to find the p-value to test whether $\beta_{\text{Age}} = 0$ using the observed t-value you computed in Question #6. Show your work/syntax.
8. Create a plot of the probability density function of the t-distribution you used in Question #7. Also shade the cumulative density associated with the p-value under this distribution. (Hint: This should look similar to Figure 4.3 in the *Probability Distributions* unit notes.)

9. Find the absolute value of the quantile of the t-distribution you used in Question #7 that is associated with the 2.5th percentile.

10. Use the value you computed in Question #9 to compute the 95% confidence interval for age. SHow your work.