Assignment 01

Using Probability Distributions

This goal of this assignment is to give you experience using probability distributions in a regression analysis. Create an R Markdown document to respond to each of the questions below. Turn in a printed version of the knitted file. This assignment is worth 8 points. Please adhere to the following guidelines for further formatting your assignment:

- All graphics should be set to an appropriate aspect ratio and sized in the knitted document so that they do not take up more room than necessary. They should also have an appropriate caption. Learn how to do this in a code chunk using knitr syntax.
- Any typed mathematics (equations, matrices, vectors, etc.) should be appropriately typeset within the document using Markdown's display equations. See here for some examples of how mathematics can be typeset in R 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.

For each question, specify the question number using a level-2 (or smaller) header.

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.

Use the *beauty.csv* data set to fit a regression model that uses professor's beauty, (`btystdave`), professor's age (`age`), and professor's sex (`female`) to explain variation in course ratings (`avgeval`). Do not trim this model; leave all predictors in the model whether they are statistically significant or not.

Part I: Model-Level Inference

1. Partition the total sum of squared error and degrees of freedom into model and error components.
2. Use the partititioning 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 `age` predictor.

6. Compute the t-statistic for the age predictor. Show your work.
7. Use the cumulative density functionality in R to find the p-value to test whether $\beta_{\text{Age}} = 0$ using the observed t-value you computed in Question #6.
8. Use R to plot the PDF 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 3 in the *probability-distribution* notes.)