HW 6

1. In this exercise we will practice the permutation test method to implement the Spearman rank correlation test for independence, which is used to detect non-linear monotone dependence.

   a) Simulate bivariate data from the relationship $y = e^x + \varepsilon$ using the following R commands:

   ```
   n = 30
   set.seed(63)
   x = runif(n, 0.5, 6)
   y = exp(x) + rnorm(n, sd = 6)
   ```

   Plot *x* vs. *y* and comment whether the relationship is linear or not.

   b) Obtain both the Pearson and Spearman correlation coefficients. The Spearman rank correlation can be obtained from the function `cor` with `method = "spearman"`.

   c) Obtain 9999 permutation replications of the Spearman correlation and produce a histogram of them.

   d) Using the replicates from part c) calculate the p-value of the permutation test and state whether you reject the null hypothesis of independence or not and why.

   e) Compare the p-value from part d) to the p-value reported by the `cor.test` function. What do you notice?

2. In this exercise, we will simulate regression data, and will then use these data to perform best subset selection.

   a) Use `set.seed(437)` and the `rnorm()` function to generate a predictor *X* of length $n = 200$, as well as the errors vector *e* of length $n = 200$. Generate histograms of each to make sure the distributions look like a bell-shaped curve.

   b) Generate a response vector *Y* of length $n = 200$ according to the cubic model:

   $$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + e$$

   where β_0, β_1, β_2, and β_3 are constants of your own choice.

   c) Perform exhaustive best subset selection in order to choose the best model from the pool of possible predictors up to the power of fifteen: $X, X^2, \ldots, X^{15}$. What is the best model according to: (i) BIC, and (ii) AIC. Include plots of AIC, and BIC, as evidence which one is the lowest value. Report the coefficients of the best models obtained.

   d) Repeat part c) for the BIC and AIC criteria only, but this time using forward selection, as shown in class with the `stepAIC` function. How does your answer compare to the results in part c)?