Problem Set 2
MFE 402: Econometrics
Professor Rossi

This problem set is designed to review material on the sampling distribution of least squares.

Question 1

Write a function in `R` to simulate from a simple regression model.

This function should accept as inputs: β_0 (intercept), $\beta1$ (slope), X (a vector of values), and σ (error standard deviation). The function should return a vector of Y values.

You will need to use the `rnorm()` function to simulate from the normal distribution. Use the `vwretd` data from the `marketRf` dataset as the X vector. Choose non-zero values for the coefficients. Include your code in your answer as well as a plot of data simulated from the model.

Question 2

    a. Use the formulas for the least squares estimators to express the least squares intercept as a weighted sum (i.e., a linear combination) of the Y values in the same way as is done in the lecture notes for the slope (see Ch1, pg 70–72).
    b. Using the formula in part (a), derive $\mathbb{E}[b_0]$ and $\text{Var}(b_0)$ and express these quantities in terms of the true parameters and the x values.

Question 3

Assume $Y = \beta_0 + \beta_1 X + \varepsilon$ with $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ and let $\beta_0 = 2$, $\beta_1 = 0.6$, and $\sigma^2 = 2$. Use the functions defined in the course notes to simulate the sampling distribution of the intercept and verify that the formulas derived in part 2(b) are correct. That is, (1) use a very large numer of simulations (10,000) to simulate the sampling distribution of b_0, (2) evaluate the theoretical formulas for $\mathbb{E}[b_0]$ and $\text{Var}(b_0)$, and then (3) compare the simulated and theoretical values.

Question 4

Hypothesis tests.

Fit a regression of the Vanguard 500 Index Fund returns (VFIAX) on the `vwretd` series. Test the following hypothesis, each time reporting the t-statistic, the p-value, and your decision (accept or reject the null):

    a. $H_0 : \beta_1 = 1$ at the 0.05 level of significance
    b. $H_0 : \beta_0 = 0$ at the 0.10 level of significance

Question 5

Standard errors and p-values.

    a. What is a standard error? How is a standard error different from the standard deviation? How does the standard error capture sampling error?

    b. A friend is working as an inventment analyst and comes to you with some output from some statistical method that you've never heard of. The friend tells you that the output has both parameter estimates and standard errors. The friend then asks, "how do I interpret and use the standard errors?" What do you say to your friend to help even though you don't even know what model is involved?

    c. Repeat part (b) for a test statistic and the p-value.

Question 6

Conditional Distributions.

Hint: use the fitted regression of `HCX` on `vwretd` to answer this question:

    a. Compute an estimate of the conditional mean of the Vanguard HCX fund given that the market is up by 5%.

    b. Compute an estimate of the conditional standard deviation of the HCX given that the market is up by 10%.