Econometrics: Problem Set 1

1. Suppose that (Y_i, X_i) satisfy our three main OLS assumptions. A random sample of size n=250 is taken and gives us:

$$\hat{Y} = \underset{(3.1)}{5.4} + \underset{(1.5)}{3.2X}, \quad R^2 = 0.26, \quad SER = 6.2$$

a. Test $H_0 : \beta_1 = 0$ vs. $H_A : \beta_1 \neq 0$ at the 5% level.
b. Construct a 95% confidence interval for β_1.
c. Suppose you learned that Y and X were independent. Would you be surprised? Explain.
d. Suppose Y and X are independent and many sample os size n=250 are drawn, regressions estimated, and (a) and (b) answered. In what fraction of the samples would H_0 from (a) be rejected? In what fraction of samples would the value $\beta_1 = 0$ be included in the confidence interval in (b)?

2. Use the data set 'cps_2009.csv' for this. First run a regression of wage on age. Make sure you turn in a copy of your script file and the R results for each part.

a. Is the coefficient on age statistically significant at the 5% level? What about the 1%?
b. Construct a 95% CI for the coefficient.
c. Repeat (a) using only people with high school diplomas (years of education = 12)
d. Repeat (a) using only people with a college degree (years of education >= 16)
e. Is the effect of age different for those with a college degree than a high school diploma?
Hint: since the subgroups are independent the SE for the difference in coefficients $(\hat{\beta}_{a,1} - \hat{\beta}_{b,1})$ is equal to $\sqrt{[SE(\hat{\beta}_{a,1})]^2 + [SE(\hat{\beta}_{b,1})]^2}$.

3*. Consider the model: $Y_i = \beta X_i + u_i$, where u_i and X_i satisfy the key OLS assumptions. Let $\overline{\beta}$ be an estimator of β that is constructed as $\overline{\beta} = \overline{Y}/\overline{X}$, where $\overline{Y}$ and $\overline{X}$ are sample means.

a. Show that $\overline{\beta}$ is a linear function of $Y_1, Y_2, ..., Y_n$.
(In other words show it can be written as: $\sum_{i=1}^{n} a_i Y_i$ for some a_is.
b. Show that $\overline{\beta}$ is conditionally unbiased, i.e. $E(\overline{\beta}|X) = \beta$
Hint: First start with the definition of $\overline{\beta}$, then replace the definition of Y and separate parts. Now use the definition of the sample mean and the linearity of the expectation function along with the fact that $E(u_i|X) = 0$.