

Assignment 2

Michael Cai

February 16, 2016

1 Problem One

(a) *Interpret the estimated intercept and slope coefficient. How would you interpret R^2 ?*

The intercept indicates that when the ratio of the U.S. consumer price index to the Canadian consumer price index equals 0 the exchange rate of the Canadian dollar to the U.S. dollar (CAD/USD) will be -0.902 CAD/USD.

The slope coefficient indicates that for a 1 unit increase in the ratio of the U.S. CPI to the Canadian CPI there is a corresponding 2.430 unit increase in the exchange rate CAD/USD.

The R^2 indicates that 0.440 or 44.4 percent of the variation in the exchange rate of CAD/USD can be explained by the ratio of the U.S. CPI to the Canadian CPI.

(b) *Does the positive value of the estimated slope coefficient make economic sense? What is the underlying economic theory?*

Yes, it does make sense because the CPI, which indicates inflation/price level, and the exchange rate are endogenously related variables. If you were to hold all other factors constant for the two currencies, then an increase in X, which is an increase in U.S. CPI/Canadian CPI, means that inflation in the U.S. is outpacing inflation in Canada. If this is the case then the U.S. dollar is depreciating relative to the Canadian dollar, and thus this is met by a corresponding increase in Y, or the exchange rate CAD/USD.

The underlying economic theory for this concept is purchasing power parity, which states that the exchange rate of two currencies must equal the ratio of the cost of a basket of goods in those two currencies.

(c) *Suppose we were to redefine X as the ratio of the Canadian CPI to the U.S. CPI. Would that change the sign of the coefficient of X? Why?*

If you redefine X as the ratio of Canadian CPI/U.S. CPI then an increasing X would imply a relative depreciation of the Canadian dollar to the U.S. dollar, which would be met by a decreasing Y, as the exchange rate of CAD/USD would be decreasing as the Canadian dollar decreases in value with respect to the U.S. dollar.

Therefore, there would be a change in the sign of the coefficient of X because upon flipping

the ratio you are effectively flipping the relation.

2 Problem Two

Consider the model $Y_i = \beta_0 + \beta_1 X_i + u_i$ where X and Y are as defined in question 1. Suppose you are told that the standard error of $\hat{\beta}_1$ is 0.096. Based on this and the estimates presented in question 1:

a) What is the p-value of a test of the hypothesis $H_0 : \beta_1 = 1$ against $H_1 : \beta_1 > 1$?

First we must calculate the t-statistic, which is defined as $t = \frac{\hat{\beta}_1 - \beta_{1,0}}{s.e.(\hat{\beta}_1)}$.

$$t = \frac{2.430 - 2}{0.096} = 4.479$$

Then we calculate the p-value from the normal distribution based on this t-statistic.

For a t-statistic of 4.479 the corresponding p-value is $1 - \Phi(4.479)$ which is < 0.00001 .

b) What is the p-value of a test of the hypothesis $H_0 : \beta_1 = 2$ against $H_1 : \beta_1 \neq 2$?

The t-statistic is calculated the same way as in part a), but the difference is that instead of computing it was $1 - \Phi(x)$ you have to double it to account for both tails. Therefore the p-value is equal to $2(1 - \Phi(4.479))$ which is still very small at < 0.00002 .

c) Yes, they are different because you must account for both tails of the distribution when calculating a two-tailed test.

d) To construct a 95 percent confidence interval, we must calculate the interval of $[\hat{\beta}_1 - 1.96(s.e.\hat{\beta}_1), \hat{\beta}_1 + 1.96(s.e.\hat{\beta}_1)]$.

$$[\hat{\beta}_1 - 1.96(s.e.\hat{\beta}_1), \hat{\beta}_1 + 1.96(s.e.\hat{\beta}_1)] = [4.479 - 1.96(0.096), 4.479 + 1.96(0.096)] \\ = [4.29084, 4.66716] \text{ is the 95 percent confidence interval for } \beta_1.$$

3 Problem Three

a) Interpret the estimated slope coefficient.

The slope coefficient of $\hat{\beta}_1 = 13.9$ is difference in the average test score between the smaller classes and the regular classes.

b) Interpret the R^2 .

The R^2 value of 0.01 indicates that 1 percent of the variation in test scores is explained by the variation the class size.

c) Is the estimated effect of the class size on test score statistically significant? Perform a test at the 5 percent level.

The first step to calculate statistical significance of a variable is to formulate hypotheses:

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

Next we need to calculate the t-statistic.

$$t = \frac{13.9 - 0}{2.5} = 5.56$$

Then we need to calculate the p-value from the t-statistic.

$$\text{p-value} = 2\Phi(5.56) \approx 0.$$

Because the p-value of 0 is less than 0.05, we can reject the H_0 and say that the effect of class size on test score is statistically significant with 95 percent confidence.

d) Construct a 95 percent confidence interval for the effect of SmallClass on TestScore.

$$[\hat{\beta}_1 - 1.96(\text{s.e.}\hat{\beta}_1), \hat{\beta}_1 + 1.96(\text{s.e.}\hat{\beta}_1)] = [13.9 - 1.96(2.5), 13.9 + 1.96(2.5)] \\ = [9.15, 18.65]$$

e) Do you think the regression errors are plausibly homoskedastic?

Yes, because the variance for a small class versus a large class in terms of test score looks like it would be the same.

4 Problem 4

a) Write the sample regression function.

$$\widehat{\text{TestScore}}_i = 931.9 - 13.9 * \text{LargeClass}_i$$

The reason why this is the new regression function is that if the LargeClass_i variable is equal to 0, i.e. when there is a small class, then the average test score value is 931.9 (or from the previous regression model, $918 + 13.9$ when the SmallClass_i variable is equal to 1).

b) Would the R^2 you obtained be different from the R^2 in question 3? Explain.

No, because the variables LargeClass_i and SmallClass_i have a dual relationship, where each variable contains exactly the same amount of information as the other. One could think of LargeClass_i and SmallClass_i as being the reverse of each other, and thus their explanatory power is the same. Consequently, this makes their R^2 values the same.

5 Problem 5

a) Construct 95 percent confidence intervals for β_0 and β_1 .

$$\hat{\beta}_0 = 623.6 \text{ and the s.e.}\hat{\beta}_0 = 7.72.$$

$$\hat{\beta}_1 = .005749 \text{ and the s.e.}\hat{\beta}_1 = .001443.$$

Therefore the 95 percent confidence interval for both betas, which is defined as, $[\beta - 1.96(\text{s.e.}\beta), \beta + 1.96(\text{s.e.}\beta)]$, is:

For $\hat{\beta}_0$: [608.469, 638.731]
 For $\hat{\beta}_1$: [.00292, .00858]

b) A local congressman claimed that increasing per-student expenditure by 1000 dollars would increase average test scores by at least 3 points. Test the congressman's claim. You may use a 5 percent level of significance.

To test this claim, we must see if a 1000 unit increase in β_1 will produce at least a 3 point increase in average test scores at the 5 percent significance level.

The model that we estimated is:

$TS_i = 623.6 + .005749Expenditure_i$ where $s.e.\beta_0 = 80.783$ and the $s.e.\beta_1 = .001443$.

Therefore the $\hat{\beta}_1$ that we are testing against is $1000 * .005749 = 5.749$ with a standard error of $1000 * .001443 = 1.443$.

We formulate our hypotheses:

$$H_0 : \beta_1 < 3$$

$$H_1 : \beta_1 \geq 3$$

Now we calculate the t-statistic:

$$t = \frac{\hat{\beta}_1 - \beta_{1,0}}{s.e.\hat{\beta}_1}$$

$$t = \frac{5.749 - 3}{1.443} = 1.905$$

Then we calculate the p-value from the normal distribution given this t-statistic.

$$p = \Phi(1.905) = .0568$$

Because this is greater than our $\alpha = 0.05$ we cannot reject the H_0 . Therefore the congressman's claim is not statistically significant.

c) Redo parts a) - b) using heteroskedasticity-consistent standard errors.

The confidence intervals for β_0 and β_1 :

$\hat{\beta}_0 = 617.87$ and the $s.e.\hat{\beta}_0 = 10.215$.

$\hat{\beta}_1 = .0057488$ and the $s.e.\hat{\beta}_1 = .0016428$.

Therefore the 95 percent confidence interval for both betas, which is defined as, $[\beta - 1.96(s.e.\beta), \beta + 1.96(s.e.\beta)]$, is:

For $\hat{\beta}_0$: [597.8486, 637.8914]

For $\hat{\beta}_1$: [.002528912, .008968688]

Testing the congressman's claim again we have the same hypotheses:

$$H_0 : \beta_1 < 3$$

$$H_1 : \beta_1 \geq 3$$

Now we calculate the t-statistic based on the values $\hat{\beta}_1$ and $s.e.\hat{\beta}_1 * 1000$:

$$t = \frac{\hat{\beta}_1 - \beta_{1,0}}{s.e.\hat{\beta}_1}$$

$$t = \frac{5.7488 - 3}{1.6428} = 1.67324$$

Seeing as this t-statistic is even lower (more moderate) than our previous one, the p-

value for this t-statistic will also be larger than the previous one, and thus changing to heteroskedasticity-consistent errors does not change the answer that the claim is not statistically significant.