

Second Midterm Exam
Economics 410
Thurs., April 2, 2009

Show All Work. Only partial credit will be given for correct answers if we can not figure out how they were derived. Note that we have not put equal value on all problems. Note as well that the t-table and F-table are provided at the back of the exam.

Points:

Problem 1:	40
Problem 2:	20
Problem 3:	20
Problem 4:	20
<hr/>	
Total:	100

Problem 1: Consider running the regression

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i.$$

You get the results

	Estimate	Standard Error
$\widehat{\beta}_0$	2.21	0.98
$\widehat{\beta}_1$	0.45	0.34
$\widehat{\beta}_2$	-0.91	0.04
$\widehat{\beta}_3$	1.89	1.21

You also know that

$$\begin{aligned} \text{cov}(\widehat{\beta}_1, \widehat{\beta}_2) &= 0.01 \\ \text{cov}(\widehat{\beta}_2, \widehat{\beta}_3) &= -0.02 \\ \text{cov}(\widehat{\beta}_1, \widehat{\beta}_3) &= 0.25 \end{aligned}$$

Suppose the sample size is 10,000.

a) Calculate a 95% confidence interval for β_0 .

The formula for a 95% confidence interval with a large number of degrees of freedom is:

$$\begin{aligned} & \left(\widehat{\beta}_0 - 1.96 \text{se} \left(\widehat{\beta}_0 \right), \widehat{\beta}_0 + 1.96 \text{se} \left(\widehat{\beta}_0 \right) \right) \\ &= (2.21 - 1.96 \times 0.98, 2.21 + 1.96 \times 0.98) \\ &= (0.289, 4.131) \end{aligned}$$

b) Calculate a 90% confidence interval for β_3 .

The formula for a 90% confidence interval with a large number of degrees of freedom is:

$$\begin{aligned} & \left(\widehat{\beta}_3 - 1.645 \text{se} \left(\widehat{\beta}_3 \right), \widehat{\beta}_3 + 1.645 \text{se} \left(\widehat{\beta}_3 \right) \right) \\ &= (1.89 - 1.645 \times 1.21, 1.89 + 1.645 \times 1.21) \\ &= (-0.100, 3.88) \end{aligned}$$

c) Test the null hypothesis: $H_0 : \beta_2 = 0$ against a two sided alternative at the 5% level.

$$\begin{aligned} T &= \frac{\widehat{\beta}_2 - 0}{\text{se} \left(\widehat{\beta}_2 \right)} \\ &= \frac{-0.91}{0.04} \\ &= 22.74 \end{aligned}$$

The critical value is 1.96 so we reject the null.

- d) Test the null hypothesis: $H_0 : \beta_1 = 0$
at the 5% level where the alternative is $\beta_1 = 5$.

$$\begin{aligned} T &= \frac{\hat{\beta}_1 - 0}{se(\hat{\beta}_1)} \\ &= \frac{0.45}{0.34} \\ &= 1.32 \end{aligned}$$

Since it is a 5% test and we want one sided, the critical value is 1.645. We fail to reject.

- e) Test the null hypothesis: $H_0 : \beta_3 = 1.0$ with the alternative $\beta_3 = 0$ at the 1% level.

$$\begin{aligned} T &= \frac{\hat{\beta}_3 - 1}{se(\hat{\beta}_3)} \\ &= \frac{1.89 - 1}{1.21} \\ &= 0.74 \end{aligned}$$

The critical value here is -2.326 so we fail to reject.

- f) Test the null hypothesis: $H_0 : \beta_2 = \beta_3$ with the alternative $\beta_2 \neq \beta_3$ at the 5% level.

$$\begin{aligned} T &= \frac{\hat{\beta}_3 - \hat{\beta}_2}{se(\hat{\beta}_3 - \hat{\beta}_2)} \\ &= \frac{\hat{\beta}_3 - \hat{\beta}_2}{\sqrt{se(\hat{\beta}_3)^2 + se(\hat{\beta}_2)^2 - 2cov(\hat{\beta}_3, \hat{\beta}_2)}} \\ &= \frac{1.89 - 0.91}{\sqrt{(1.21)^2 + (0.04)^2 + 2 \times 0.02}} \\ &= \frac{2.8}{1.227} \\ &= 2.281 \end{aligned}$$

This is greater than 1.96 so we reject the null.

- g) Construct a 95% confidence interval for

$$\theta = 2 + \beta_1 - \beta_3.$$

In this case

$$\begin{aligned}\hat{\theta} &= 2 + \hat{\beta}_1 - \hat{\beta}_3 \\ &= 2 + 0.45 - 1.89 \\ &= 0.56\end{aligned}$$

$$\begin{aligned}se(\hat{\theta}) &= \sqrt{\text{var}(2 + \hat{\beta}_1 - \hat{\beta}_3)} \\ &= \sqrt{se(\hat{\beta}_1)^2 + se(\hat{\beta}_3)^2 - 2\text{cov}(\hat{\beta}_1, \hat{\beta}_3)} \\ &= \sqrt{(0.34)^2 + (1.21)^2 - 2 \times 0.25} \\ &= 1.04\end{aligned}$$

The formula for a 95% confidence interval with a large number of degrees of freedom is:

$$\begin{aligned}&(\hat{\theta} - 1.96se(\hat{\theta}), \hat{\theta} + 1.96se(\hat{\theta})) \\ &= (0.56 - 1.96 \times 1.04, 0.56 + 1.96 \times 1.04) \\ &= (-1.48, 2.60)\end{aligned}$$

Problem 2: Suppose you are interested in knowing the causal effect of teacher experience (E_i) on test scores (T_i) of kids:

$$T_i = \beta_0 + \beta_1 E_i + u_i$$

where T_i is the score out of 100 on a test. Think about this in terms of a causal model making the appropriate assumptions.

You are reported evidence of this from one of two studies:

- In the first case the estimate of $\hat{\beta}_1$ would be 5 with a standard error of 3 from a sample size of 500.
- In the second case the estimate of $\hat{\beta}_1$ would be .05 with a standard error of .03 from a sample size of 5000.

a) Suppose your null hypothesis is that teacher experience is irrelevant. State this as a null hypothesis and test the null for both cases (two sided test, 5% size). What do you conclude from this?

Case 1:

$$\begin{aligned} T &= \frac{\hat{\beta}_1 - 0}{se(\hat{\beta}_1)} \\ &= \frac{5}{3} \\ &= 1.67 \end{aligned}$$

Case 2:

$$\begin{aligned} T &= \frac{\hat{\beta}_1 - 0}{se(\hat{\beta}_1)} \\ &= \frac{0.05}{0.03} \\ &= 1.67 \end{aligned}$$

The critical value for a two sided test is 1.96 so I fail to reject the null hypothesis in both of these cases at the 95% level. Since the T-statistics are the same the evidence against the null is similar. That is, the p-values will be identical.

b) Now construct 95% confidence intervals from each of the two studies.

Case 1:

$$\begin{aligned} & \left(\hat{\beta}_1 - 1.96se(\hat{\beta}_1), \hat{\beta}_1 + 1.96se(\hat{\beta}_1) \right) \\ &= (5 - 1.96 \times 3, 5 + 1.96 \times 3) \\ &= (-0.88, 10.88) \end{aligned}$$

Case 2:

$$\begin{aligned} & \left(\hat{\beta}_1 - 1.96se\left(\hat{\beta}_1\right), \hat{\beta}_1 + 1.96se\left(\hat{\beta}_1\right) \right) \\ &= (0.05 - 1.96 \times 0.03, 0.05 + 1.96 \times 0.03) \\ &= (-0.0088, 0.109) \end{aligned}$$

- c) Based on your answer to b), what do you conclude about the causal effect from each of the two studies? In particular, how would your conclusions differ across the two studies.

The main difference from the two studies was that the second is much more informative. The confidence interval in Case 1 is much much wider and includes many more values. If I look at case 2 I am confident that a year of teacher experience leads to an increase in a test score of 0.1%. This seems like a small number to me so from case 2 it looks like the effect of teacher experience is small. However, in case 1 I am confident that the effect is no larger than 10.88% but this seems like a huge effect. The confidence interval also includes zero though, so I learn very little about this from the first case.

Problem 3: Suppose you are interested in the effects of the price of a cell phone on the number that are bought. Consider the regression model of the demand:

$$C_i = \beta_0 + \beta_1 P_i + \beta_2 R_i + u_i$$

and assume that the classical regression model holds so that this is causal. Here

C_i : Number of Cell Phones Bought
 P_i : “Sticker” price of cell phone
 R_i : Rebate you get for buying phone

According to economic theory all that matters is net price (i.e. the sticker price - the rebate).

a) State the economic theory implication as a null hypothesis that you can test.

Economic theory says that all that matters is net price. This means if I increase the price by a dollar and increase the rebate by a dollar, I would expect no effect. The effect of increase price by a dollar and Rebate by a dollar is $\beta_1 + \beta_2$. So the null hypothesis that all that matters is net price can be stated as

$$H_0 : \beta_1 + \beta_2 = 0$$

b) Explain how to “trick stata” into running a regression that will allow you to test the null hypothesis directly (that is directly from the stata output).

To trick stata I will follow something similar to what we did in class: Define $\theta = \beta_1 + \beta_2$, then

$$\begin{aligned} C_i &= \beta_0 + \beta_1 P_i + \beta_2 R_i + u_i \\ &= \beta_0 + \beta_1 P_i + (\theta - \beta_1) R_i + u_i \\ &= \beta_0 + \beta_1 (P_i - R_i) + \theta R_i + u_i \end{aligned}$$

So to test the null hypothesis that $\beta_1 + \beta_2 = \theta = 0$, we can run a regression of C_i on net price ($P_i - R_i$) and the rebate and then just test whether the coefficient on the rebate is zero.

Problem 4: Consider the regression model

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + u_i.$$

Suppose you have 95 observations.

Explain how you would test the null hypothesis

$$\begin{aligned} H_0 : \beta_1 &= -\beta_4 \\ \beta_2 &= 2 + \beta_3 \end{aligned}$$

at the 5% level.

We need to run two models. The model above which will be the unrestricted model. Calculate the R^2 of that model (and call it R_u^2) We also run a restricted model which we can write as

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + u_i \\ &= \beta_0 - \beta_4 X_{1i} + (2 + \beta_3) X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + u_i \\ Y_i - 2X_{2i} &= \beta_0 + \beta_3(X_{2i} + X_{3i}) + \beta_4(X_{4i} - X_{1i}) + u_i \end{aligned}$$

That is for the restricted regression I run a regression of $(Y_i - 2X_{2i})$ on $(X_{2i} + X_{3i})$ and $(X_{4i} - X_{1i})$ and get the R^2 from this regression (call it R_r^2). We then construct the F statistic as

$$F = \frac{(R_u^2 - R_r^2)/2}{(1 - R_u^2)/(95 - 4 - 1)}$$

and then compare it to the critical value which in this case is 3.10.