6104641060
Suchaya Chuldilok

1. Which of the following can cause the usual OLS *t* statistics to be invalid (that is, not to have *t* distributions under H_0)?

i. Heteroskedasticity.

Heteroskedasticity can cause the usual OLS t statistics to be invalid; assumption of homoscadasticity is violated.

ii. A sample correlation coefficient of .95 between two independent variables that are in the model.

No, it requires just the coefficient not to be "1".

iii. Omitting an important explanatory variable.

Yes, violate the 4^{th} assumption $E(u|x)=0$ if the omitted variable is correlated with the independent variables in the model.

2. Consider an equation to explain salaries of CEOs in terms of annual firm sales, return on equity (*roe*, in percentage form), and return on the firm's stock (*ros*, in percentage form):

$$\log(salary) = \beta_0 + \beta_1 \log(sales) + \beta_2 roe + \beta_3 ros + u.$$

i. In terms of the model parameters, state the null hypothesis that, after controlling for *sales* and *roe*, *ros* has no effect on CEO salary. State the alternative that better stock market performance increases a CEO's salary.

$H_0: \beta_3 = 0$

$H_1: \beta_3 > 1$

ii. Using the data in CEOSAL1, the following equation was obtained by OLS:

$$\widehat{\log(salary)} = \underset{(.32)}{4.32} + \underset{(.035)}{.280} \log(sales) + \underset{(.0041)}{.0174}\, roe + \underset{(.00054)}{.00024}\, ros$$

$$n = 209, R^2 = .283.$$

By what percentage is *salary* predicted to increase if *ros* increases by 50 points? Does *ros* have a practically large effect on *salary*?

The proportionate effect on $\widehat{\log(salary)}$ is $(0.00024)(50) = 0.012$.
Since this is a log-linear regression we multiply this by 100 in order to obtain the % change, which is 1.2%. Therefore, a 50-point ceteris paribus increase in ros is predicted to increase salary by only 1.2%. Practically speaking, this is a very small effect for such a large change in ros.

iii. Test the null hypothesis that *ros* has no effect on *salary* against the alternative that *ros* has a positive effect. Carry out the test at the 10% significance level.

We first need to obtain the 10% critical value for one-tailed test. Since n is large I will use the approximation at $df = \infty$. This comes out to 1.282. We now need to calculate the t-statistic for this test, which is

$t = \dfrac{(\hat{\beta}_3 - 0)}{(s.e.(\hat{\beta}_3))} = \dfrac{0.00024}{0.00054} = 0.44 < 1.282$. We failed to reject the null hypothesis.

iv. Would you include *ros* in a final model explaining CEO compensation in terms of firm performance? Explain.

Based on the sample, the estimated ros coefficient appears to be different from zero only because of sampling variation. On the other hand, if ros is very correlated with the other explanatory variables then omitting it from the regression may still bias the estimates.

Chapter Review : Computer Excercises

C6. Use the data in WAGE2 for this exercise.

i. Consider the standard wage equation

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + u.$$

$H_0: \beta_2 = \beta_3$

ii. Test the null hypothesis in part (i) against a two-sided alternative, at the 5% significance level, by constructing a 95% confidence interval. What do you conclude?

C.6

```
. reg lwage educ exper tenure

      Source |       SS           df       MS      Number of obs   =       935
-------------+----------------------------------   F(3, 931)       =     56.97
       Model |  25.6953242         3  8.56510806   Prob > F        =    0.0000
    Residual |  139.960959       931  .150334005   R-squared       =    0.1551
-------------+----------------------------------   Adj R-squared   =    0.1524
       Total |  165.656283       934  .177362188   Root MSE        =    .38773

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0748638   .0065124    11.50   0.000      .062083    .0876446
       exper |   .0153285   .0033696     4.55   0.000     .0087156    .0219413
      tenure |   .0133748   .0025872     5.17   0.000     .0082974    .0184522
       _cons |   5.496696   .1105282    49.73   0.000     5.279782    5.713609
------------------------------------------------------------------------------
```

According the estimation, α is insignificant at a 95% confidence interval, hence we cannot reject the null hypothesis. Therefore, we conclude that 1 additional year of general work force experience has the same effect on log(wage).

6104641060
Suchaya Chuldilok

C8. The data set 401KSUBS contains information on net financial wealth (*nettfa*), age of the survey respondent (*age*), annual family income (*inc*), family size (*fsize*), and participation in certain pension plans for people in the United States. The wealth and income variables are both recorded in thousands of dollars. For this question, use only the data for single-person households (so $fsize = 1$).

i. How many single-person households are there in the data set?

There are 2,017 observations left

ii. Use OLS to estimate the model

$$nettfa = \beta_0 + \beta_1 inc + \beta_2 age + u,$$

and report the results using the usual format. Be sure to use only the single-person households in the sample. Interpret the slope coefficients. Are there any surprises in the slope estimates?

we interpret $\hat{\beta}_1$ as a \$1000 increase in income corresponds to a \$799 increase in net financial wealth. we interpret $\hat{\beta}_2$ as a 1 year increase in age corresponds to a \$842 increase in net financial wealth.

iii. Does the intercept from the regression in part (ii) have an interesting meaning? Explain.

$\hat{\beta}_0 = -43.04$. This is an individual's net financial wealth when their income is \$0 and their age is 0, so this is the net financial wealth of new born babies which we are unlikely to be interested in.

iv. Find the *p*-value for the test $H_0\colon \beta_2 = 1$ against $H_1\colon \beta_2 < 1$. Do you reject H_0 at the 1% significance level?

The t-statistic is $(0.843 - 1)/0.092 \approx -1.71$. Against the one-sided alternative $H_1: \beta_1 < 1$ the p-value is $P(T < -1.71) \approx 0.044$. we can reject the null hypo. at 5% significance level but not at the 1% significance level.

v. If you do a simple regression of *nettfa* on *inc*, is the estimated coefficient on *inc* much different from the estimate in part (ii)? Why or why not?

we get an estimate of $\hat{\beta}_1 = 0.821$ which is not very different from the estimate of 0.799 in the last reg. Since this is an omitted variable question we need to know the correlated btw. age and income, which is 0.039. It can explain that the coefficient does not change much.

C.8

```
. reg nettfa inc age if fsize ==1

      Source |       SS           df       MS      Number of obs   =     2,017
-------------+----------------------------------   F(2, 2014)      =    136.46
       Model |  544916.989         2  272458.495   Prob > F        =    0.0000
    Residual |  4021048.06     2,014  1996.54819   R-squared       =    0.1193
-------------+----------------------------------   Adj R-squared   =    0.1185
       Total |  4565965.05     2,016  2264.86361   Root MSE        =    44.683

------------------------------------------------------------------------------
      nettfa |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inc |   .7993167   .0597307    13.38   0.000     .6821762    .9164572
         age |   .8426563   .0920169     9.16   0.000     .6621982    1.023115
       _cons |  -43.03981   4.080393   -10.55   0.000    -51.04204   -35.03758
------------------------------------------------------------------------------

. reg nettfa inc if fsize == 1

      Source |       SS           df       MS      Number of obs   =     2,017
-------------+----------------------------------   F(1, 2015)      =    181.60
       Model |  377482.064         1  377482.064   Prob > F        =    0.0000
    Residual |  4188482.98     2,015   2078.6516   R-squared       =    0.0827
-------------+----------------------------------   Adj R-squared   =    0.0822
       Total |  4565965.05     2,016  2264.86361   Root MSE        =    45.592

------------------------------------------------------------------------------
      nettfa |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inc |   .8206815      .0609    13.48   0.000     .7012479     .940115
       _cons |  -10.57095   2.060678    -5.13   0.000    -14.61223   -6.529671
------------------------------------------------------------------------------

. corr nettfa inc age
(obs=9,275)

             |   nettfa      inc      age
-------------+---------------------------
      nettfa |   1.0000
         inc |   0.3766   1.0000
         age |   0.2039   0.1056   1.0000
```

C1. The following model can be used to study whether campaign expenditures affect election outcomes:

6104641060

Suchaya Chuldilok

$$voteA = \beta_0 + \beta_1 \log(expendA) + \beta_2 \log(expendB) + \beta_3 prtystrA + u,$$

where *voteA* is the percentage of the vote received by Candidate A, *expendA* and *expendB* are campaign expenditures by Candidates A and B, and *prtystrA* is a measure of party strength for Candidate A (the percentage of the most recent presidential vote that went to A's party).

i. What is the interpretation of β_1 ?

The elasticity of candidate A's expense. 1% of change of ExpendA is associated with β_1% change of vote A.

ii. In terms of the parameters, state the null hypothesis that a 1% increase in A's expenditures is offset by a 1% increase in B's expenditures.

$H_0: \beta_1 = -\beta_2$

iii. Estimate the given model using the data in VOTE1 and report the results in usual form. Do A's expenditures affect the outcome? What about B's expenditures? Can you use these results to test the hypothesis in part (ii)?

C.1

```
. reg voteA lexpendA lexpendB prtystrA
```

Source	SS	df	MS			
				Number of obs	=	173
				F(3, 169)	=	215.23
Model	38405.1096	3	12801.7032	Prob > F	=	0.0000
Residual	10052.1389	169	59.480112	R-squared	=	0.7926
				Adj R-squared	=	0.7889
Total	48457.2486	172	281.728189	Root MSE	=	7.7123

voteA	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lexpendA	6.083316	.38215	15.92	0.000	5.328914	6.837719
lexpendB	-6.615417	.3788203	-17.46	0.000	-7.363246	-5.867588
prtystrA	.1519574	.0620181	2.45	0.015	.0295274	.2743873
_cons	45.07893	3.926305	11.48	0.000	37.32801	52.82985

Both expense of A & B affect the outcome.

iv. Estimate a model that directly gives the *t* statistic for testing the hypothesis in part (ii). What do you conclude? (Use a two-sided alternative.)

$H_0: \beta_1 + \beta_2 = 0$

Let $\alpha = \beta_1 + \beta_2$

$H_0: \alpha = 0$

$H_1: \alpha \neq 0$

$\beta_1 = \alpha - \beta_2$

```
. test lexpendA-lexpendB = 0

 ( 1)  lexpendA - lexpendB = 0

       F(  1,   169) =  546.81
            Prob > F =    0.0000
```