6104640302

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.

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

   iii. Omitting an important explanatory variable.

(i) and (iii) generally cause the t statistic not to have t distribution under H_0. Homoskedasticity is one of the CLM assumption. An important omitted variable violates Assumption MLR 3. The CLM assumption contains no mention of the sample correlation among independent variable, except to rule out the case where the correlation is one.

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.

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*?

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.

Print Preview

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

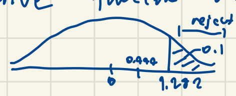
(i) $H_0: \beta_3 = 0 \qquad H_a: \beta_3 > 0$

(ii) The proportionate effect on $\widehat{salary}$ is $0.00024(50) = 0.12$

To obtain the percentage effect we multiply this by 100 : 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

(iii) The 10% critical value for one-tailed test, using df = ∞ is obtained from table G2 as 1.282. The t statistic on ros is $\frac{0.00024}{0.00054} = 0.44$ which is well below the critical value therefore we fail to reject Ho at the 10% significance level.



(iv) No, since we already test that ros does not have any effect on salary at 0.1 level of significance. Even though if we add it will raise to r^2, but it will worsen the variance, we only want to put in what really explain the model.

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

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

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 ?

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.

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)?

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.)

C i) β_1 is the point change in vote A when exped A increcre by one

C ii) The null hypothesis is H_0 $\beta_2 = -\beta_1$ which means a 2% increcse in expenditure by A and a 2% increcre in expenditure by B leaves vote A unchanged. We can equivalently write H_0 $\beta_1 + \beta_2 = 0$

C iii) $\text{Vote A} = 45.08 + 6.083 \log(\text{expend A}) - 6.615 \log(\text{expend B}) + 0.152\, \text{prytstrA} + u$

$t \text{ test } A = \frac{6.083}{0.382} = 15.924$

$t \text{ cri } A = (169, 0.05) = 1.654$

$t \text{ test } B = \frac{-6.615}{0.379} = -17.454$

$t \text{ cri } B \ (169, 0.05) = 1.654$

reject A, reject B

Both A and B expenditure have significant effect on outcome

can't use this model on hypothesis in part (ii)

(iv) gen log(expend A ÷ expend B) = log expend A − log expend B

voteA = 42.7 + 6.342 log(expend A / expend B) + 0.146 prtystrA + u

$t_{cri} = \frac{6.342}{0.271} = 23.4$

$t_{cri}(170, 0.05) \approx 1.654$

⇒ reject H_0, 1% increse in A expenditure

is not off set by 1% increace in B's expendi

C6. Use the data in WAGE2 for this exercise.

i. Consider the standard wage equation

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

State the null hypothesis that another year of general workforce experience has the same effect on $\log(wage)$ as another year of tenure with the current employer.

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?

(i) the null hypothesis is $\beta_2 = \beta_3$

(ii) let $\theta_2 = \beta_2 - \beta_3$, then we can estimate the equation in the question to obtain 95% of CI for θ_2. This turn out to be about $0.002 \pm 1.96(0.0047) = -0.0072$ to 0.112. Because zero is in this CI, θ_2 is not statistically different from 0 at the 5% level, and we fail to reject $H_0: \beta_2 = \beta_3$, at the 5% level

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?

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?

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

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?

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?

C i) There are 2017 single people

C ii) $\widehat{nettfa} = -43.04 + 0.799\, inc + 0.843\, age$

The coefficient on inc indicates that one more dollar in income is reflected in about 0.799 more cents in predicted nettfa

The coefficient on age means that if age increase by 1 the predicted nettfa is predicted to increase by \$843

C iii) The intercept = -43.04 when inc = 0 and age = 0.

C iv) The t statis is -1.71. The one side alternative $H_1: \beta_2 < 1$ the p-value is about 0.44, Therefore, we can reject $H_0: \beta_2 = 1$ at the 1% significant level

C v) The slope coefficient on inc is 0.821 whic is not different 0.799 in part C ii)

The correlation between inc and age of sample of single people is about 0.39 which help explain why simple and multiple regression estimates are not very different