

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.

ANS: YES, it's violate the assumption of homoscedasticity because OLS method includes homoscedasticity as one of its assumption, but if this assumption does not hold, that means if the data turn out to be heteroskedastic then the OLS t statistic will be invalid.

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

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

iii. Omitting an important explanatory variable.

ANS: Yes, violated the 4th assumption which is $E(u|x) = 0$ if the omitted is correlated with the independent variable in the model it will definitely cause the t statistic to be invalid.

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):

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

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.

ANS: $H_0: \beta_3 = 0$
 $H_a: \beta_3 > 0$

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

$$\text{OLS: } \widehat{\log(\text{salary})} = 4.32 + .280 \log(\text{sales}) + .0174 roe + .00024 ros$$

$$(.32) \quad (.035) \quad (.0041) \quad (.00054)$$

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

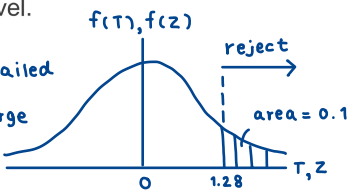
By what percentage is $salary$ predicted to increase if ros increases by 50 points? Does ros have a practically large effect on $salary$?

ANS: The effect on $\widehat{\text{salary}}$ is $0.00024(50) = 0.012$ or 1.2%. means that if ros increases by 50 points, $salary$ will predicted to increase by 1.2%. which 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.

ANS: one-tailed

n is relatively large
then we will use
Z-table.



$$t_{\hat{\beta}_3} = \frac{(\hat{\beta}_3 - \beta_3)}{s.e.(\hat{\beta}_3)} = \frac{0.00024 - 0}{0.00054} = 0.44$$

So, $t_{\hat{\beta}_3} < \text{Critical value } (0.44 < 1.28)$ means that we do not reject H_0 at 10% significance level. Therefore, ros has no effect on $Salary$.

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

ANS The estimated *ros* coefficient appears to be different from zero only because of sampling variation. On the other hand, including *ros* may not be causing any harm. It depends on how correlated it is with the other independent variable.

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

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

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 ?

ANS The elasticity of candidate A's expense. If *Expend A* change by 1%, *Vote A* will change by $\beta_1\%$.

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.

ANS $H_0: \beta_1 = -\beta_2$ or $\beta_1 + \beta_2 = 0$

$H_a: \beta_1 \neq -\beta_2$ or $\beta_1 + \beta_2 \neq 0$

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)?

ANS

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

Source	SS	df	MS	Number of obs	=	173
Model	38405.1096	3	12801.7032	F(3, 169)	=	215.23
Residual	10052.1389	169	59.480112	Prob > F	=	0.0000
				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 A's expenditures and B's expenditures affect the outcome significantly. According to the estimation, 1% increase in A's expenditure would be increase *VoteA* by 6.08%. Nevertheless, 1% increase in B's expenditure would be decrease *Vote B* by -6.62%. Moreover, you cannot test the null hypothesis from part (ii) because we do not have the standard error of $\hat{\beta}_1 + \hat{\beta}_2$.

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.)

ANS $\text{vote}_A = \beta_0 + \beta_1 \log(\text{expend}_A) + \beta_2 \log(\text{expend}_B) + \beta_3 \text{Prtystr}_A + u$

Would like to test the $H_0: \beta_1 + \beta_2$

Let $\theta = \beta_1 + \beta_2$

1st) Write the t -statistic for testing H_0

$$t = \frac{(\hat{\beta}_1 - 3\hat{\beta}_2) - 1}{\text{s.e.}(\hat{\beta}_1 - 3\hat{\beta}_2)} \quad t = \frac{(\beta_1 + \beta_2) - 0}{\text{s.e.}(\beta_1 + \beta_2)}$$

2nd) Define $\theta_1 = \beta_1 + \beta_2 \gg H_0: \theta_1 = 0, H_a: \theta_1 \neq 0$

$$t = \frac{\theta_1 - 1}{\text{s.e.}(\theta_1)}$$

Let $\theta = \beta_1 + \beta_2$ and sub in the main regression and get

$$\text{vote}_A = \beta_0 + \beta_1 \log(\text{expend}_A) + \beta_2 \log(\text{expend}_B) + \beta_3 \text{Prtystr}_A + u$$

$$= \beta_0 + (\beta_1 + \beta_2 - \beta_2) \log(\text{expend}_A) + \beta_2 \log(\text{expend}_B) + \beta_3 \text{Prtystr}_A + u$$

$$= \beta_0 + \theta \log(\text{expend}_A) - \beta_2 \log(\text{expend}_A) + \beta_2 \log(\text{expend}_B) + \beta_3 \text{Prtystr}_A + u$$

$$= \beta_0 + \theta \log(\text{expend}_A) + \beta_2 (\log(\text{expend}_B) - \log(\text{expend}_A)) + \beta_3 \text{Prtystr}_A + u$$

Source	SS	df	MS	Number of obs =	173
Model	38405.1089	3	12801.703	F(3, 169) =	215.23
Residual	10052.1397	169	59.4801165	Prob > F =	0.0000
Total	48457.2486	172	281.728189	R-squared =	0.7926
				Adj R-squared =	0.7889
				Root MSE =	7.7123

votea	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lexpenda	-0.532101	.5330858	-1.00	0.320	-1.584466 .520264
exp	-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

According to the result table, it obtain that $\theta_1 = -0.532$ and $\text{s.e.}\theta_1 = 0.533$.

The t -statistic for the hypothesis in part (ii) is $\frac{-0.532}{0.533} \approx -1$. Therefore,

we will not reject the null hypothesis means that 1% increase in A's expenditures is offset by a 1% increase in B's expenditures.

C6. Use the data in WAGE2 for this exercise.

i. Consider the standard wage equation

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

State the null hypothesis that another year of general workforce

experience has the same effect on $\log(\text{wage})$ as another year of tenure

with the current employer.

ANS $H_0: \beta_2 = \beta_3$ or $\beta_2 - \beta_3 = 0$

$H_a: \beta_2 \neq \beta_3$ or $\beta_2 - \beta_3 \neq 0$

- 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?

```
. reg lwage educ exper tenure
```

Source	SS	df	MS	Number of obs	=	935
Model	25.6953242	3	8.56510806	F(3, 931)	=	56.97
Residual	139.960959	931	.150334005	Prob > F	=	0.0000
				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

ANS : According to the estimation table, $\alpha = 0.05$ is not significant at a 95% of confidence interval. So, we will not reject H_0 . Therefore, one additional year of general workforce experience has the same effect on $\log(\text{wage})$ as another year.

- 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?

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

Source	SS	df	MS	Number of obs	=	2,017
Model	544916.989	2	272458.495	F(2, 2014)	=	136.46
Residual	4021048.06	2,014	1996.54819	Prob > F	=	0.0000
				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

ANS According to the estimation table, the number of single-person households are 2017 observations.

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?

ANS The slope coefficients is $\hat{\beta}_1$ and $\hat{\beta}_2$. With $\hat{\beta}_1$, we can interpret that if income increase by thousands of dollar (\$1000), net finance wealth would be increase by \$799. However, with $\hat{\beta}_2$, we can interpret that if age increase by 1 year, net finance wealth would be increase by \$842.

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

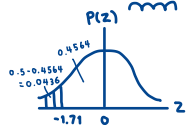
ANS $\hat{\beta}_0$ is a intercept of regression function means that when their income is 0 and their age is 0, individual's net financial wealth is \$-43.04. Therefore, this is the net financial wealth of newborn babies which it does not seems to be interesting. We can conclude that there is none of them even close to these values in the appropriate population.

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

ANS From the output that we obtain in part (ii), we can conclude that t-statistic

$$\text{is } \frac{\hat{\beta}_2 - \beta_2}{S.E.(\hat{\beta}_2)} = \frac{0.843 - 1}{0.092} \approx -1.71 \text{ and the p-value for age is } P(T < -1.71) \approx 0.0436$$

Therefore, we will reject H_0 at 1% significant level because p-value is less than 1% significant level (0.01).



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?

ANS The estimated coefficient on $\hat{\beta}_1 = 0.821$ which it is not different much from the estimate of 0.799 in the last regression. Since this is omitted variable, we need to know the correlated between age and income which we can find to be just only 0.039 means that coefficient does not change much.