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.

iii) Omitting an important explanatory variable that means your the regression model will have a bias. because the important explanatory variable has impact on 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.

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* $= 0$ against the alternative that *ros* has a positive effect. > 0 Carry out the test at the 10% significance level.

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

i) $H_0: \beta_3 = 0$ $\quad t = \dfrac{\beta_3 - \beta_3}{s.e. \cdot \beta_3}$

$H_a: \beta_3 > 0$

ii) $\dfrac{\partial \log(salary)}{\partial \log(ros)} = \dfrac{d\,(0.00024)\,ros}{d \log(ros)} - 0.00024 \dfrac{\cancel{d}\, ros}{\frac{1}{ros}\, \cancel{d} ros}$

$= .00024 \times (ros) = 0.00024(50) = \mathbf{1.2\,\%}$

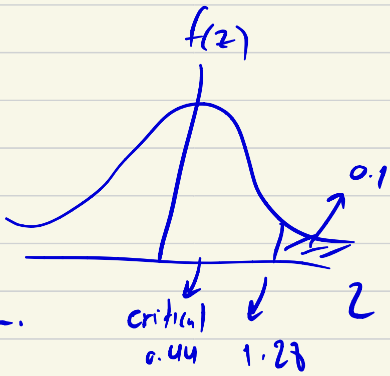
if the firm's stock increase 1%, the salary will change. 1.2 %. No, there's just small effect.

iii) $\alpha = 0.1$, $n = 209$ → z-table.

$H_0: \beta_3 = 0$ $\quad z = \dfrac{\hat{\beta}_3 - \bar{\beta}_3}{s.e.\, \hat{\beta}_3} = \dfrac{0.00024}{0.00054} = 0.44$

$H_a: \beta_3 > 0$

$z_{critical} = 0.44$, $z_{0.1} = 1.28$



so z critical doesn't fall in the rejected region. that mean ros has no effect on salary

iv) I will includ ros in final model because t-test proof that ros has no effect so if I put ros in model. it has nothing happen.

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.)

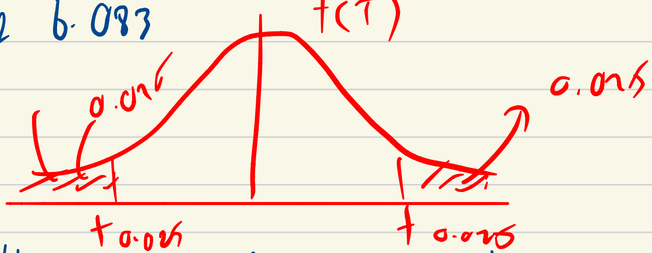
```
. regress voteA lexpendA lexpendB prtystr
```

Source	SS	df	MS			
Model	38405.1096	3	12801.7032	Number of obs	=	173
Residual	10052.1389	169	59.480112	F(3, 169)	=	215.23
				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

VotA = 45.0789 + 6.083316 lexpendA − 6.615417 lexpendB + 0.15196 prtystrA + u.

i) when the percentage of the vote receive by expendA increase by 1 % then Vote A will increase by 6.083



ii) $H_0: \beta_1 = -\beta_2 \rightarrow \beta_1+\beta_2=0$

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

iii) from P>|t| = 0.00 of both lexpend A and lexpend B that means the t critical of both variable fall in rejected region when $\bar{\beta}_j = 0$, it means A and B's expenditure has impact on outcome. I think we can't use ii) test for finding lexpendA and lexpend B have impact on outcome.

C6. Use the data in WAGE2 for this exercise.

i. Consider the standard wage equation

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

Print Preview

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) $H_0: \beta_2 = \beta_3$

$H_a: \beta_2 \neq \beta_3$

$\downarrow$

$H_0: \beta_2 - \beta_3 = 0$

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

ii) let $\hat{\beta}_2 - \hat{\beta}_3 = \hat{\theta}_1$

so $H_0: \hat{\theta}_1 = 0$

$H_a: \hat{\theta}_1 \neq 0$

$t = \hat{\theta}_1 / s.e.\ \hat{\theta}_1$

if rearrange $\hat{\beta}_2 = \hat{\theta}_1 + \hat{\beta}_3$ and then substitute $\hat{\beta}_2 = \hat{\theta}_1 + \hat{\beta}_3$ in the regression model.

$$\log(wage) = \beta_0 + \beta_1 educ + (\hat{\theta}_1 + \hat{\beta}_3) exper + \beta_3 tenure + u$$

$$= \beta_0 + \beta_1 educ + \hat{\theta}_1 exper + \hat{\beta}_3 exper + \beta_3 tenure + u$$

$$\log(wage) = \beta_0 + \beta_1 educ + \hat{\theta}_1 exper + \beta_3 (exper + tenure) + u$$

. regress lwage educ exper expertenure

Source	SS	df	MS
Model	25.6953242	3	8.56510806
Residual	139.960959	931	.150334005
Total	165.656283	934	.177362188

Number of obs	=	935
F(3, 931)	=	56.97
Prob > F	=	0.0000
R-squared	=	0.1551
Adj R-squared	=	0.1524
Root MSE	=	.38773

lwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
educ	.0748638	.0065124	11.50	0.000	.062083	.0876446
exper	$\hat{\theta}_1$ .0019537	.0047434	0.41	0.681	-.0073554	.0112627
expertenure	.0133748	.0025872	5.17	0.000	.0082974	.0184522
_cons	5.496696	.1105282	49.73	0.000	5.279782	5.713609

$t_{critical} = 0.41$ less than 1.96

that means the null hypotheses isn't rejected because t critical doesn't fall in rejected region. so both of variables has no the same impact.

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 $H_0: \beta_2 = 1$ against $H_1: \beta_2 < 1$. 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?

. regress nettfa inc age if fsize == 1

Source	SS	df	MS
Model	544916.989	2	272458.495
Residual	4021048.06	2,014	1996.54819
Total	4565965.05	2,016	2264.86361

Number of obs	=	2,017
F(2, 2014)	=	136.46
Prob > F	=	0.0000
R-squared	=	0.1193
Adj R-squared	=	0.1185
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

. sum if fsize == 1

Variable	Obs	Mean	Std. Dev.	Min	Max
e401k	2,017	.3604363	.4802461	0	1
inc	2,017	29.44618	16.67356	10.008	143.067
marr	2,017	.0183441	.1342256	0	1
male	2,017	.5418939	.4983654	0	1
age	2,017	39.27814	10.82328	25	64
fsize	2,017	1	0	1	1
nettfa	2,017	13.59498	47.59058	-143.5	1134.098
p401k	2,017	.2429351	.4289625	0	1
pira	2,017	.2141795	.4103536	0	1
incsq	2,017	1144.947	1581.761	100.1601	20468.17
agesq	2,017	1659.857	922.5799	625	4096

i) from STATA, there are 2,017 observation

ii) from STATA (regress nettfa inc age if fsize = 1)

$\beta_0 = -43.03981$, $\beta_1 = 0.79931$, $\beta_2 = 0.84265$

$nettfa = -43.03981 + 0.79931\, inc + 0.84265\, age + u.$

the slope of family income and slope of age means if the family income or age increas by 1 unit, the net financial wealth will increase by 0.79931 or 0.84265 and if the family income and age are zero, the net financial wealth will equal to -43.03981.

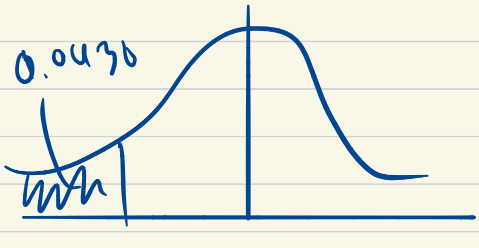
the impact of age greater than the family income in the single-person households in the sample.

iii) yes, if the family income and age equal to zero, the net financial wealth will equal to -43.03981 or loss in net financial weath equals to 43.03981

iv) $\left.\begin{array}{l} H_0: \beta_2 = 1 \\ H_a: \beta_2 < 1 \end{array}\right\}$ 1 tailed test we use z-table $f(z)$

$$z = \frac{\hat{\beta}_2 - \beta_2}{S.E.\,\hat{\beta}_2} = \frac{0.843 - 1}{0.092} = -1.71$$

0.0436



-1.71

So we don't reject the null hypotheses because 0.0436 > 0.01 at 1% signifcand level.

v.)

. regress nettfa inc if fsize == 1

Source	SS	df	MS
Model	377482.064	1	377482.064
Residual	4188482.98	2,015	2078.6516
Total	4565965.05	2,016	2264.86361

Number of obs = 2,017
F(1, 2015) = 181.60
Prob > F = 0.0000
R-squared = 0.0827
Adj R-squared = 0.0822
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

it differents about 0.02 of both coefficients. and when the variable has only the family income impact net financial wealth more than the multiple regression that may be because when more variable the relationship between variables effect to the net financial wealth.