6109640252
Jirathorn C.

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?

```
. 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

There are 2,017 single-person households 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?

```
. regress 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
intercept _cons	-43.03981	4.080393	-10.55	0.000	-51.04204	-35.03758

- When inc rises by 1 unit, nettfa will increase by 0.7993 units.
- when age rises by 1 unit, nettfa will increase by 0.843 units
- Both inc and age have positive impact on nettfa

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

The intercept from the regression in part (ii) can be interpreted that when other factors are zero, we can expect the net financial wealth to be -43.03981

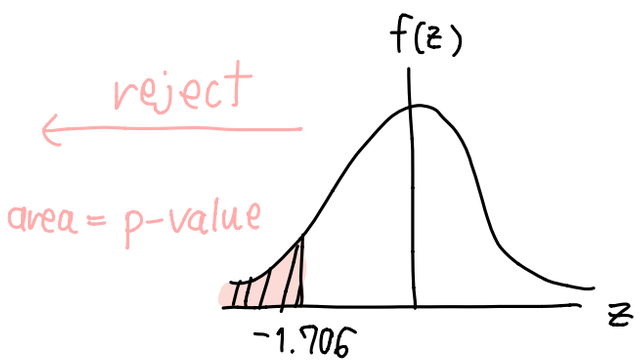
iv. Find the *p*-value for the test $\mathbf{H_0}: \beta_2 = 1$ against $\mathbf{H_1}: \beta_2 < 1$. Do you reject $\mathbf{H_0}$ at the 1% significance level?

$\hat{\beta}_2 = 0.843$

$H_0 : \beta_2 = 1$, $Ha : \beta_2 < 1$

$$t = \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} = \frac{0.843 - 1}{0.092}$$

$$= -1.706$$



area = 0.5 − 0.4608

= 0.0392

since p-value > 0.01 , so we donot reject H_0

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 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

when comparing with (ii), we will see that the coef. value in this problem (v) is greater because (v) has only one explanatory variable, inc. In part (ii), we include more explanatory variables, so they will share the impact on the explained variables (less value of coef.).