ECN 102, Spring 2020

Midterm 2 Review
Regression

Variables `sales` is how many sales (in units) a firm makes; variable `newspaper` is thousands of dollars a firm spends for newspaper advertising.

```
. regress sales newspaper
```

Source	SS	df	MS	Number of obs	=	200
				F(1, 198)	=	10.89
Model	282344204	1	282344204	Prob > F	=	0.0011
Residual	5.1348e+09	198	25933356.3	R-squared	=	0.0521
				Adj R-squared	=	0.0473
Total	5.4171e+09	199	27221853	Root MSE	=	5092.5

sales	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
newspaper	54.6931	16.57572	3.30	0.001	22.00548	87.38071
_cons	12351.41	621.4202	19.88	0.000	11125.96	13576.86

```
t_198,.005 = 2.6008873
 t_198,.01 = 2.3453283
t_198,.025 = 1.9720175
 t_198,.05 = 1.6525858
 t_198,.10 = 1.2858418
```

Question 1

How do sales change when newspaper advertising expenditure increases by one thousand dollars?

```
. regress sales newspaper
```

Source	SS	df	MS		
Model	282344204	1	282344204	Number of obs	= 200
Residual	5.1348e+09	198	25933356.3	F(1, 198)	= 10.89
				Prob > F	= 0.0011
Total	5.4171e+09	199	27221853	R-squared	= 0.0521
				Adj R-squared	= 0.0473
				Root MSE	= 5092.5

sales	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
newspaper	54.6931	16.57572	3.30	0.001	22.00548	87.38071
_cons	12351.41	621.4202	19.88	0.000	11125.96	13576.86

The slope coefficient is $b_2 = 54.6931$. An increase in variable `newspaper` by 1 unit (which is one thousand dollars) is associated with 54.6931 more units sold.

Question 2

The claim is made that sales are not associated with newspaper advertising. Test this claim at significance level 0.05. State clearly the null and alternative hypothesis and your conclusion.

```
. regress sales newspaper
```

Source	SS	df	MS	Number of obs	=	200
				F(1, 198)	=	10.89
Model	282344204	1	282344204	Prob > F	=	0.0011
Residual	5.1348e+09	198	25933356.3	R-squared	=	0.0521
				Adj R-squared	=	0.0473
Total	5.4171e+09	199	27221853	Root MSE	=	5092.5

sales	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
newspaper	54.6931	16.57572	3.30	0.001	22.00548	87.38071
_cons	12351.41	621.4202	19.88	0.000	11125.96	13576.86

"Not associated with" is code for zero slope. The null hypothesis is $H_0 : \beta_2 = 0$ against $H_a : \beta_2 \neq 0$. *Stata tests this automatically.* Notice the p-value for `newspaper` is $0.001 < 0.05$, so we reject the null and conclude that the association between newspaper and sales is statistically significant at level 0.05.

Question 3

The claim is made that saless change by more than 30 units when advertising expenditure increases by one thousand dollars. Test at level 0.05. State clearly the null and alternative hypotheses and your conclusion.

```
. regress sales newspaper
```

Source	SS	df	MS		
Model	282344204	1	282344204	Number of obs	= 200
Residual	5.1348e+09	198	25933356.3	F(1, 198)	= 10.89
				Prob > F	= 0.0011
Total	5.4171e+09	199	27221853	R-squared	= 0.0521
				Adj R-squared	= 0.0473
				Root MSE	= 5092.5

sales	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
newspaper	54.6931	16.57572	3.30	0.001	22.00548	87.38071
_cons	12351.41	621.4202	19.88	0.000	11125.96	13576.86

```
t_198,.005 = 2.6008873
 t_198,.01 = 2.3453283
t_198,.025 = 1.9720175
 t_198,.05 = 1.6525858
 t_198,.10 = 1.2858418
```

The claim is $H_a : \beta_2 > 30$ and the null is $H_0 : \beta_2 \leq 30$. The t-statistic is

$$t = \frac{54.9631 - 30}{16.57572} = 1.506,$$

and the critical value is $t_{200-2,0.05} = 1.6525858$, so fail to reject the null.

Question 4

Suppose \$100,000 is spent on newspaper advertising. What level of sales do we expect?

```
. regress sales newspaper
```

Source	SS	df	MS		
Model	282344204	1	282344204	Number of obs	= 200
Residual	5.1348e+09	198	25933356.3	F(1, 198)	= 10.89
				Prob > F	= 0.0011
				R-squared	= 0.0521
				Adj R-squared	= 0.0473
Total	5.4171e+09	199	27221853	Root MSE	= 5092.5

sales	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
newspaper	54.6931	16.57572	3.30	0.001	22.00548	87.38071
_cons	12351.41	621.4202	19.88	0.000	11125.96	13576.86

The estimated regression line is $\hat{y} = 12351.41 + 54.6931x$. Each unit of x is \$1000, so \$100,000 is equivalent to $x = 100$. Plug this into the estimated regression line to get a predicted value of

$$\hat{y} = 12351.41 + 54.6931(100) = 17820.72 \text{ units sold.}$$