

Exercise 1.1

I have observations on systolic blood pressure (SBP) and age for a sample of 30 adults. A linear regression using age to predict SBP was conducted; Stata output is below.

```
. regress sbp age
```

Source	SS	df	MS	Number of obs = 30		
Model	6394.02269	1	6394.02269	F(1, 28) = 21.33		
Residual	8393.44398	28	299.765856	Prob > F = 0.0001		
Total	14787.4667	29	509.912644	R-squared = 0.4324		
				Adj R-squared = 0.4121		
				Root MSE = 17.314		

sbp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	.9708704	.2102157	4.62	0.000	.5402629	1.401478
_cons	98.71472	10.00047	9.87	0.000	78.22969	119.1997

(a) Interpret the intercept estimate. Is it meaningful?

(b) Interpret the slope estimate.

(c) What is the estimated mean SBP for a person who is 30 years old?

(d) What is the estimated difference in mean SBP for a person who is 45 years old compared to a person who is 30 years old? (And clearly state which mean is higher.)

(e) Is there evidence of a significant association between age and SBP? Cite specific evidence from the output (i.e., a p-value).

(f) What is the coefficient of determination for this model? Write a one sentence interpretation of this quantity.

Exercise 1.2

Information on 74 automobiles was collected (in 1978) to study the relationship between gas mileage (mpg) and various features of the cars. In particular, we are interested in the relationship between mileage and weight of the car, measured in pounds. A linear regression produced the (incomplete) Stata output below.

```
. regress mpg pounds
```

Source	SS	df	MS	Number of obs	=	74
-----+-----				F(1, 72)	=	134.62
Model	1591.9902	1	1591.9902	Prob > F	=	
Residual	851.469256	72	11.8259619	R-squared	=	
-----+-----				Adj R-squared	=	0.6467
Total	2443.45946	73	33.4720474	Root MSE	=	3.4389

mpg	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
-----+-----						
pounds	-.0060087	.0005179				
_cons	39.44028	1.614003	24.44	0.000	36.22283	42.65774
-----+-----						

(a) Write the estimated regression equation.

(b) Interpret the number 39.44 in the above output.

(c) Interpret the number -0.0060087 in the above output.

(d) What is the estimated change in MPG for a 1 U.S. ton increase in car weight? (Note: 1 U.S. ton = 2000 pounds)

(e) Test whether there is a significant effect of weight on mileage.

(f) Calculate a 95% confidence interval for β_1 . Do you expect 0 to be in the interval?

(g) The R^2 value is missing from the output. Calculate it.

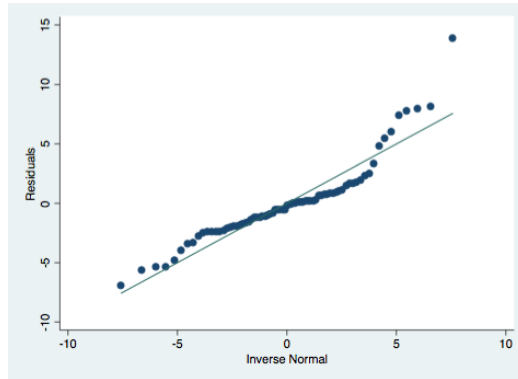
(h) Calculate the correlation between MPG and pounds.

(i) If I convert the weight of the car from pounds to kilograms and rerun the regression model (i.e., use weight in kilograms to predict MPG), will the slope estimate change? Will the R^2 change?

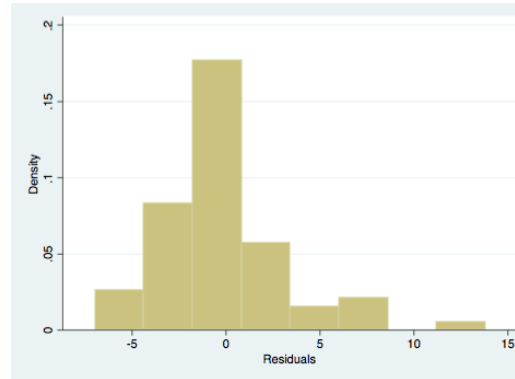
(j) The p-value for the overall F-test is missing from the output. What is it?

(k) Three plots generated after running the regression model are shown below (labeled A-C). For each plot, state the name of the plot (or what is being plotted), which assumption(s) of the model can be checked with the plot, and whether or not you see any problems with the assumptions.

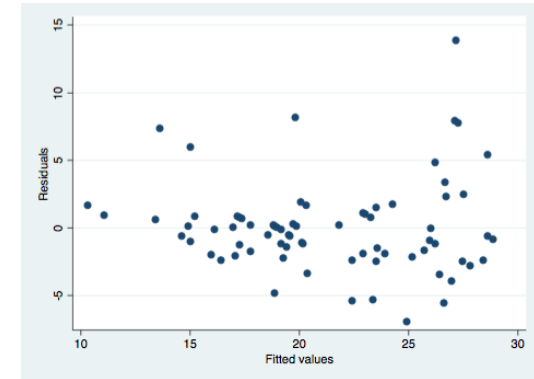
(A)



(B)



(C)



Exercise 1.3

An economist is interested in the relationship between money flowing to stock mutual funds and money flowing into bond mutual funds. She collects data on the net new money flow into stocks and bonds for each year from 1985 to 2000 (in billions of dollars) and adjusts for inflation. She then uses Stata to run the regression model: $E(BONDS) = \beta_0 + \beta_1 \times STOCKS$

```
. regress bonds stocks
```

Source	SS	df	MS	Number of obs	=	16
				F(1, 14)	=	1.60
Model	5749.11511	1	5749.11511	Prob > F	=	-----
Residual	50200.7429	14	3585.76735	R-squared	=	-----
				Adj R-squared	=	0.0387
Total	55949.858	15	3729.99053	Root MSE	=	59.881

bonds	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
stocks	-.1962223	.1549669	-1.27	-----	-.5285933	.1361486
_cons	53.40959	22.9925	2.32	0.036	4.09557	102.7236

(a) The p-value for the test of $H_0 : \beta_1 = 0$ is missing from the output. Based only on the confidence interval, what do you know about the missing p-value? (Hint: would you reject or fail to reject H_0 ?)

(b) The R^2 value is also missing. Calculate it.