EconHw3

Trevor Freeland

April 23, 2018

```r
library(tidyverse)
library(readxl)
library(stringr)
library(pander)
```

```r
table6 <- read_xls("~/R/Econometrics/Tables/Table7_6.xls")
table6 <- table6[-(1:10),]
colnames(table6) <- table6[1,]
table6 <- table6[-1,]
table6[] <- sapply(table6[], as.numeric)
table9 <- read_xls("~/R/Econometrics/Tables/Table7_9.xls")
table9 <- table9[-(1:9),]
colnames(table9) <- table9[1,]
table9 <- table9[-1,]
table9[] <- sapply(table9[], as.numeric)
table4 <- read_xls("~/R/Econometrics/Tables/Table6_4.xls")
table4 <- table4[-1,]
colnames(table4) <- table4[1,]
table4 <- table4[-1,]
table4[] <- sapply(table4[], as.numeric)
savings <- read.csv("~/R/Econometrics/EconHw3/PersonalSaving.csv")
income <- read.csv("~/R/Econometrics/EconHw3/PersonalIncome.csv")
personal <- left_join(income, savings)
```

1

A: False

B: False

C: False

D: True

2

(A)

α_2 = -2228 -> Holding all else constant, if the average wholesale price of roses went up by \$1/dozen, on average the number of roses sold would decrease by 2227.

α_3 = 1251 -> Holding all else constant, if the average wholesale price of carnations went up by \$1/dozen, on average the number of roses sold would increase by 1251.

α_4 = 6.3 -> Holding all else constant, if the average weekly family disposable income grew by \$1/week, on average the number of roses sold would increase by 6.

$\alpha_5 = $ -197 -> Holding all else constant, on average the quantity of roses sold in a quarter decreased by 197 roses per quarter starting in 1971-III.

See table below for full model summary.

```
rose.lm <- lm(Y~X2 + X3 + X4 + X5, data = table6)
pander(summary(rose.lm))
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	10816	5988	1.806	0.0983
X2	-2228	920.5	-2.42	0.034
X3	1251	1157	1.081	0.3027
X4	6.283	30.62	0.2052	0.8412
X5	-197.4	101.6	-1.944	0.07796

Table 2: Fitting linear model: Y ~ X2 + X3 + X4 + X5

Observations	Residual Std. Error	R^2	Adjusted R^2
16	969.9	0.8347	0.7746

(B)

$\beta_2 = $ -1.17 -> Holding all else constant, if the average wholesale price of roses went up by 1% per dozen, on average the number of roses sold would decrease by 1.17%.

$\beta_3 = $ .74 -> Holding all else constant, if the average wholesale price of carnations went up by 1% per dozen, on average the number of roses sold would increase by .74%.

$\beta_4 = $ 1.15 -> Holding all else constant, if the average weekly family disposable income grew by 1% per week, on average the number of roses sold would increase by 1.15%.

$\beta_5 = $ -.03 -> Holding all else constant, on average the quantity of roses sold in a quarter decreased by 3% roses per quarter starting in 1971-III.

See table below for full model summary.

```
rose.loglm <- lm(log(Y)~log(X2) + log(X3) + log(X4) + X5, data = table6)
pander(summary(rose.loglm))
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.572	4.695	0.7608	0.4628
log(X2)	-1.171	0.4883	-2.397	0.03539
log(X3)	0.7379	0.6529	1.13	0.2824
log(X4)	1.153	0.902	1.279	0.2274
X5	-0.03011	0.01642	-1.834	0.09384

Table 4: Fitting linear model: $\log(Y) \sim \log(X2) + \log(X3) + \log(X4) + X5$

Observations	Residual Std. Error	R^2	Adjusted R^2
16	0.1607	0.7988	0.7256

2

(C)

Own-Price: We would expect as the price of the product increased, demand would decrease. (Holding everything else constant)

Cross-Price: We would expect as the price of a competitor's product increases, the demand for our product would increase. (Holding everything else constant)

Income Elasticity: We would expect as our buyers had more money to spend, that the demand for our product would increase. (Holding everything else constant)

Our estimates in our model concur with our expectations.

(D)

Own-Price: 11.5%

Cross-Price: 5.6%

Income Elasticity: .02%

See code for calculations.

```
own.price <- ((10816-(10816-2228))/(10816+(10816-2228)))/(1/1)
cross.price <- (1/(10816+10186+1251)) * (1251)/(1)
income <- ((6)/(10816+10816+6))/(1/1)
```

(E)

I would choose to use the log-log model for our analysis. Especially in this setting, it seems we are very interested in Own-Price/Cross-Price/Income-Elasticity, which are all far easier to find, understand and interpret using a log-log model. Also knowing the scales of both of the variables are extremely large, typically you would use a log-log model instead of a linear model.

(F)

In order to test that the demand in own-price unit elastic we need to see if for every% increase in cost we see an equal percent decrease in sales. What this means for our regression model is that we want to test to see if the estimate for $\log(X2) = -1$. So lets make a test stat:

$$t = \frac{-1.17 - (-1)}{.49} = -.35$$

Even before calculating the actual p-value we know that with a t stat under 1 standard deviation away from the hypothesized value that we won't be able to reject the null hypothesis. The p-value we get when testing our t-stat is .37, which is to large to be any evidence that our null hypothesis is wrong, and so we can not reject the null hypothesis that the demand is own-price unit elastic.

```
t.stat <- (-1.17 + 1)/.49
p <- pt(t.stat, 11)
```

3

(A)

I would chose the **5th** demand function. The 4th and 5th demand functions have very similar information, but there is likely to be correlation between X4 and X5, so if we combine them into one variable we won't have to worry about high correlation between them.

(B)

The coefficient of lnX2 would be interpreted as the income elasticity of demand for chicken.

The coefficient of lnX3 would be interpreted as the own-price elasticity of demand for chicken.

(C)

Model (2) is not accounting for the effect of beef prices in the consumption of chickens whereas model (4) does take into account the price of beef.

(D)

There is likely to be correlation between the prices of beef and pork, and so having them both in the model could lead to issues.

(E)

I would prefer model (5) because it gets rid of the potential correlation and colinearity effects of having both the individual prices of beef and the individual prices of pork in the model.

(F)

They are substitute products for chicken. The reason we know this is because as the price of beef and pork go up, hodling everything else constant, the consumption of chicken rises. This is behavior indicative of substitute products.

To verify this thought, I ran the regression of chicken consumption soley against prices of pork and beef individually and both estimates were posotive, indicating an increase in pork or beef price leads to an increase in chicken consumption, implying that beef and pork are substitutes for chicken.

```
testmodel <- lm(log(Y)~log(X4), data = table9)
coef(testmodel)
```

```
(Intercept)      log(X4)
  1.6400039    0.4564348
```

```
testmodel2 <- lm(log(Y)~log(X5), data = table9)
coef(testmodel2)
```

```
(Intercept)      log(X5)
  1.4728259    0.4611974
```

(G)

Our model: $Y = \beta_0 + \beta_1(\text{DisposableIncome}) + \beta_2(\text{PriceChicken}) + \beta_3(\text{PriceChickenSubs}) + \epsilon,\ \epsilon \sim N(0,\sigma)$

We will focus on the interpretations of β_1, β_2, β_3. For full model summary and other estimates see table below.

$\beta_1 = .48$ -> Holding all else constant, if the average real disposable income per capita increased by 1%, on average the number of chickens consumed per capita (lbs) would increase by .41%.

$\beta_2 = -.35$ -> Holding all else constant, if the average real retail price of chicken per lb increased by 1%, on average the number of chickens consumed per capita (lbs) would decrease by .35%.

$\beta_3 = -.06$ -> Holding all else constant, if the average composite real price of chicken substitues per lb increased by 1%, on average the number of chickens consumed per capita (lbs) would decrease by .06%.

```
model5 <- lm(log(Y)~log(X2) + log(X3) + log(X6), data = table9)
pander(summary(model5))
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2.03	0.1187	17.1	5.36e-13
log(X2)	0.4813	0.06819	7.058	1.022e-06
log(X3)	-0.3506	0.07939	-4.416	0.0002965
log(X6)	-0.06104	0.13	-0.4696	0.644

Table 6: Fitting linear model: $\log(Y) \sim \log(X2) + \log(X3) + \log(X6)$

Observations	Residual Std. Error	R^2	Adjusted R^2
23	0.02834	0.9803	0.9772

(H)

Here are the difference gamma and beta values:

$\gamma_2 = .41$

$\gamma_3 = -.44$

$\beta_2 = .48$

$\beta_3 = -.35$

Because we are only including the price of pork and are not accounting for the price of beef at all, we are changing our income and own-price elasticity estimates. We are saying that the same change in price is leading to a larger decrease in sales in model 2 and we are saying the the same change in income is leading to a smaller increase in chicken sales in model 2. The differences are not that large her, but have incorrect other variables could potentially greatly change the estimates of our effects of interests.

See below for model 2 summary.

```
model2 <- lm(log(Y)~log(X2) + log(X3) + log(X4), data = table9)
pander(summary(model2))
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2.125	0.1379	15.42	3.401e-12
log(X2)	0.4059	0.04479	9.063	2.507e-08

	Estimate	Std. Error	t value	Pr(>|t|)
log(X3)	-0.4388	0.08333	-5.266	4.404e-05
log(X4)	0.1067	0.08784	1.214	0.2395

Table 8: Fitting linear model: $\log(Y) \sim \log(X2) + \log(X3) + \log(X4)$

Observations	Residual Std. Error	R^2	Adjusted R^2
23	0.02746	0.9815	0.9786

4

Running the simulation 10 times, and running it 1000 times we get on average estimates that are slightly larger than the true population estimates. I doubled checked by changing the seeds multiple times and I never got average estimates that were equal to or smaller than the true population estimates, always slightly larger.

See code below for simulations and results.

```
set.seed(123456)
N <- 64
sims <- 10
intercept <- numeric(sims)
beta1 <- numeric(sims)
beta2 <- numeric(sims)

for (i in 1:sims){
  rand <- rnorm(N, 0, 42)
  Y <- table4$CM - rand
  model <- lm(Y~PGNP+FLR, data = table4)
  intercept[i] <- coef(model)[1]
  beta1[i] <- coef(model)[2]
  beta2[i] <- coef(model)[3]
}
mean(intercept)
```

```
[1] 264.3491
```

```
mean(beta1)
```

```
[1] -0.005228368
```

```
mean(beta2)
```

```
[1] -2.294658
```

```
set.seed(1256)
N <- 64
sims <- 1000
intercept <- numeric(sims)
beta1 <- numeric(sims)
beta2 <- numeric(sims)

for (i in 1:sims){
  rand <- rnorm(N, 0, 42)
```

```
  Y <- table4$CM - rand
  model <- lm(Y~PGNP+FLR, data = table4)
  intercept[i] <- coef(model)[1]
  beta1[i] <- coef(model)[2]
  beta2[i] <- coef(model)[3]
}
mean(intercept)
```

```
[1] 263.4697
```
```
mean(beta1)
```

```
[1] -0.005619054
```
```
mean(beta2)
```

```
[1] -2.232172
```

5

Unfortunately I wasn't able to complete problem 5, I ran out of time. I was also running into the fact that my data is very different from the data used in the textbook and so it didn't seem to really be working with everything we were being asked to use.

(A)

PSAVE -> Save

MEPAINUSA672N -> Personal Income

(B)

These are definitely not the same plots as the ones in the book but this is what I got with the data that I had.
```
personal$year <- as.numeric(substring(personal$DATE,1,4))
personal$Income <- personal$MEPAINUSA672N
par(mfrow = c(1,2))
line1 <- lm(PSAVE~Income, data = filter(personal, year %in% c(1974:1981)))
line2 <- lm(PSAVE~Income, data = filter(personal, year %in% c(1982:1995)))
plot(PSAVE~Income, data = filter(personal, year %in% c(1974:1981)), ylab="Savings", main="1974-1981")
abline(line1)
plot(PSAVE~Income, data = filter(personal, year %in% c(1982:1995)), ylab="Savings", main = "1982-1995")
abline(line2)
```

## 1974–1981



## 1982–1995



(C)

```
par(mfrow = c(1,2))
line3 <- lm(PSAVE~Income, data = filter(personal, year > 1995) )
line4 <- lm(PSAVE~Income, data = personal )
plot(PSAVE~Income, data = personal, ylab="Savings", main="1996-2016")
abline(line3)
plot(PSAVE~Income, data = personal, ylab="Savings", main = "1974-2016")
abline(line4)
```

1996–2016   **1974–2016**
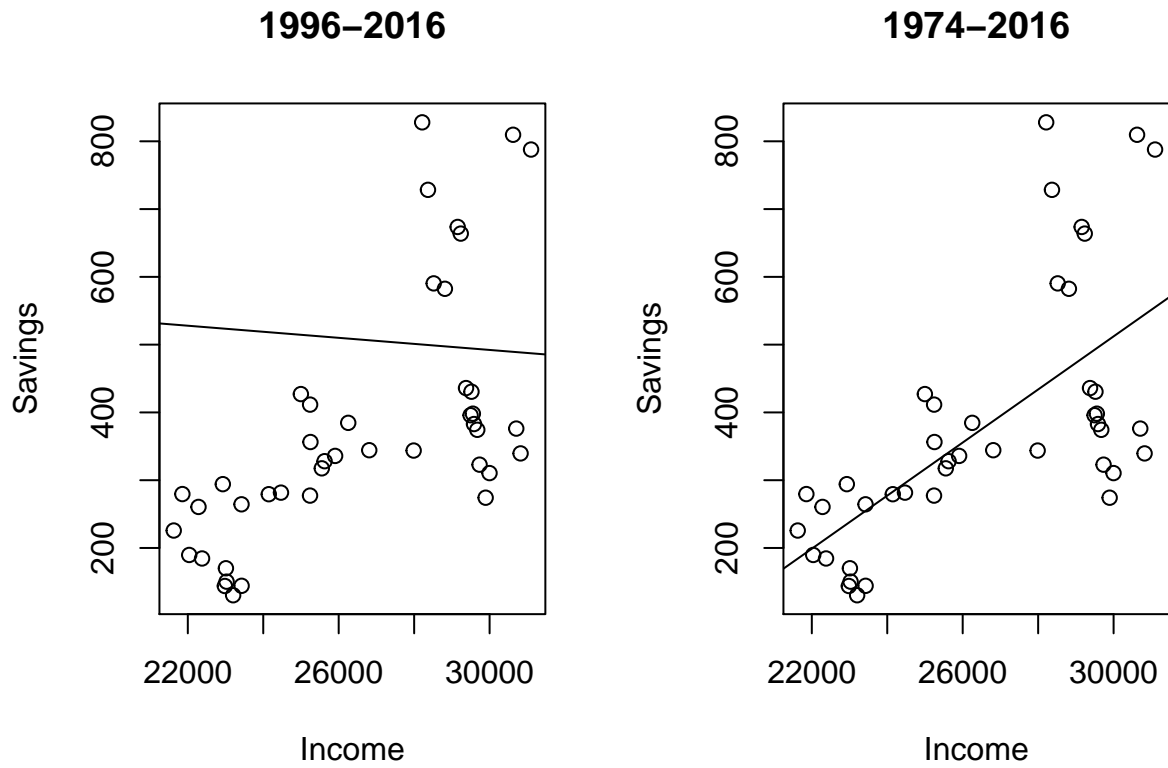
(D)

```
line5 <- lm(PSAVE~Income, data = filter(personal, year < 1995))
pander(summary(line1))
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1237	162.9	7.596	0.0002709
Income	-0.04711	0.007169	-6.571	0.0005957

Table 10: Fitting linear model: PSAVE ~ Income

Observations	Residual Std. Error	R^2	Adjusted R^2
8	11.9	0.878	0.8577

```
pander(summary(line2))
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-280.3	221.8	-1.264	0.2303
Income	0.02454	0.009033	2.717	0.01872

9

Table 12: Fitting linear model: PSAVE ~ Income

Observations	Residual Std. Error	R^2	Adjusted R^2
14	45.15	0.3808	0.3292

```
pander(summary(line5))
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-730.9	258.2	-2.831	0.01067
Income	0.04172	0.01086	3.843	0.001097

Table 14: Fitting linear model: PSAVE ~ Income

Observations	Residual Std. Error	R^2	Adjusted R^2
21	67.13	0.4373	0.4077

(E)

(F)

(G)

Challenge Problem