1. Exercise 6.7, page 230

The gross income and tax paid by a cross section of 30 companies in 1988 and 1989 is given in the Table 6.2.

1988		1989	
Income	**Tax**	**Income**	**Tax**
9.215	1.643	9.518	2.125
2.047	0.413	2.068	0.565
9.989	1.752	9.992	2.221
8.321	1.408	8.515	1.905
4.588	0.838	4.389	0.943
4.736	0.748	5.015	1.051
3.596	0.577	3.811	0.819
4.83	0.752	4.939	1.015
4.508	0.761	4.539	1.096
7.506	1.331	7.806	1.654
4.052	0.548	4.583	0.836
6.015	1.121	6.345	1.602
7.775	1.316	8.227	1.877
3.105	0.503	3.129	0.698
2.215	0.514	2.691	0.246
5.676	1.057	6.015	1.146
5.554	0.942	5.702	1.221
5.36	0.803	5.15	0.92
10.394	1.902	10.579	2.15
3.473	0.513	3.341	0.57
4.022	0.868	4.4	0.917
6.119	1.067	6.682	1.157
3.362	0.559	3.487	0.678
7.203	1.318	7.557	1.637
3.874	0.58	3.929	0.515
7.259	1.138	7.636	1.721
2.13	0.414	2.169	0.433
7.528	1.331	7.862	1.461
9.578	1.662	9.997	2.166
2.015	0.351	2.259	0.447

(a)

Use these data to estimate the relationship

$$\text{tax}_t = \beta_1 + \beta_2 \text{income}_t + e_t$$

for each of the years 1988 and 1989.

For 1988:

$$\hat{\text{tax}}_{1988} = -0.018 + 0.1763\hat{\text{income}}_{1988}$$

For 1989:

$$\hat{\text{tax}}_{1989} = -0.1085 + 0.2266\hat{\text{income}}_{1989}$$

(b)

Give the interpretations of the two estimates of β_2

In the year 1988, for any additional unit increase of income (in millions of dollars), the company will be taxed an additional 0.1736 million dollars, ceteris paribus.

In the year 1989, for any additional unit increase of income (in millions of dollars), the company will be taxed an additional 0.2266 million dollars, ceteris paribus.

(c)

Find the average income for each year and predict the tax paid for each average income. Compare the average and marginal tax rates.

$$\begin{aligned} \text{Average income}_{1988} &= \frac{\sum \text{Income}}{N} && (1)\\ &= 5.5348 && (2) \end{aligned}$$

$$\begin{aligned} \text{Average income}_{1989} &= \frac{\sum \text{Income}}{N} && (1)\\ &= 5.7444 && (2) \end{aligned}$$

Predicted tax paid for average income in 1988:

$$\begin{aligned} \hat{\text{tax}}_{1988} &= -0.018 + 0.1763(5.5348) && (1)\\ &= 0.9577 && (2) \end{aligned}$$

Predicted tax paid for average income in 1989:

$$\begin{aligned} \hat{\text{tax}}_{1989} &= -0.018 + 0.1763(5.7444) && (1)\\ &= 1.1931 && (2) \end{aligned}$$

Marginal tax rates for 1988:

$$\begin{aligned} M\text{tax}_{1988} &= \frac{d\hat{\text{tax}}_{1988}}{d\text{income}_{1988}} && (1)\\ &= \frac{d}{d\text{income}_{1988}}(-0.018 + 0.1763\hat{\text{income}}_{1988}) && (2)\\ &= 0.1763 && (3) \end{aligned}$$

Marginal tax rates for 1989:

$$\begin{aligned} M\text{tax}_{1989} &= \frac{d\hat{\text{tax}}_{1989}}{d\text{income}_{1989}} && (1)\\ &= \frac{d}{d\text{income}_{1989}}(-0.1085 + 0.2266\hat{\text{income}}_{1989}) && (2)\\ &= 0.2266 && (3) \end{aligned}$$

Average tax rates for 1988:

$$\begin{aligned} &= \frac{\sum \text{Tax}_{1988}}{\sum \text{Income}_{1988}} && (1)\\ &= \frac{28.73}{166.045} && (2)\\ &= 0.173 && (3) \end{aligned}$$

Average tax rates for 1989:

$$= \frac{\sum \text{Tax}_{1989}}{\sum \text{Income}_{1989}} \quad (1)$$
$$= \frac{35.792}{172.332} \quad (2)$$
$$= 0.2077 \quad (3)$$

For 1988, the marginal tax rate was 0.1763 and the average tax rate was 0.173. For 1989, the marginal tax rate was 0.2266 and the average tax rate was 0.2077.

(d)

Consider the estimate for β_2 for 1989 and the corresponding estimated variance $v\hat{a}r(b_2)$. Pretend that $v\hat{a}r(b_2)$ is the same as the true variance $var(b_2)$, and assume that b_2 is normally distributed. Find the probability that the sampling error $[b_2 - \beta_2]$ is (i) less than 0.04 and (ii) less than 0.1.

$$var(\hat{b}) = \begin{bmatrix} 0.0040 & -0.0006 \\ -0.0006 & 0.0001 \end{bmatrix}$$

So $var(b_1) = 0.0040$ and $var(b_2) = 0.0001$

(i)

To find $P(b_2 - \beta_2 < 0.04)$:

$$Z = \frac{0.04}{\sqrt{b_2}} \quad (1)$$
$$= \frac{0.04}{\sqrt{0.0001}} \quad (2)$$
$$= 4 \quad (3)$$

$P(b_2 - \beta_2 < 0.04) \approx 1$

(ii)

To find $P(b_2 - \beta_2 < 0.01)$:

$$Z = \frac{0.01}{\sqrt{b_2}} \quad (1)$$
$$= \frac{0.01}{\sqrt{0.0001}} \quad (2)$$
$$= 1 \quad (3)$$

$P(b_2 - \beta_2 < 0.01) = 0.8413$

(e)

Pool the observations from the two years of data and use the resulting 60 observations to estimate one tax-income relationship. Compare the estimates for β_1 and β_2, and the estimated variances $v\hat{a}r(b_1)$ and $v\hat{a}r(b_2)$, with those from the separate equations. What implicit assumptions are you making when you pool the two sets of observations?

The relationship between tax and income for both 1988 and 1989 can be given as:

$$\hat{\text{tax}}_{both} = -0.0734 + 0.2037\hat{\text{income}}_{both}$$

When pooling data sets together, we are assuming that the elements we are estimating are homogeneous, that the underlying components are the same, with the same parameter values. In other words, we are

2. Exercise 6.12, page 232

Consider the linear model $y_t = x_{t1}\beta_1 + x_{t2}\beta_2 + e_t$, where the e_t are independent random variables with mean zero and variance σ^2. Suppose we have only the following observations:

y_t	x_{t1}	x_{t2}
2	1	0
-1	0	1
4	1	1
0	1	-1

Note that this a special model without an intercept or a constant term.

(a)

Define y, X, and β such that the model can be written in matrix algebra form.

$$y = \begin{bmatrix} 2 \\ -1 \\ 4 \\ 0 \end{bmatrix}, X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

Thus, in matrix algebra form, this can be written as: $y = X\beta$

(b)

Find $X'X$, $X'y$, and $y'y$.

$$X'X = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1+0+1+1 & 0+0+1-1 \\ 0+0+1-1 & 0+1+1+1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

$$X'y = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 2+0+4+0 \\ 0-1+4+0 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$y'y = \begin{bmatrix} 2 & -1 & 4 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 4+1+16+0 \end{bmatrix} = \begin{bmatrix} 21 \end{bmatrix}$$

(c)

Find b_1 and b_2.

$$b = (X'X)^{-1}X'y \tag{1}$$

$$= \frac{1}{9}\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 6 \\ 3 \end{bmatrix} \tag{2}$$

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \tag{3}$$

(d)

Compute $\hat{e} = y - Xb = y - \hat{y}$ and use the estimator $\dfrac{\hat{e}'\hat{e}}{T-2}$ to compute an estimate of the variance σ^2.

$$Xb = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \tag{1}$$
$$= \begin{bmatrix} 2 \\ 1 \\ 3 \\ 1 \end{bmatrix} \tag{2}$$

$$\hat{e} = y - Xb \tag{1}$$
$$= \begin{bmatrix} 2 \\ -1 \\ 4 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \\ 3 \\ 1 \end{bmatrix} \tag{2}$$
$$\hat{e} = \begin{bmatrix} 0 \\ -2 \\ 1 \\ -1 \end{bmatrix} \tag{3}$$

$$\sigma^2 = \frac{\hat{e}'\hat{e}}{T-2} \tag{1}$$
$$= \frac{\begin{bmatrix} 0 & -2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \\ 1 \\ -1 \end{bmatrix}}{4-2} \tag{2}$$
$$= \frac{4+1+1}{2} \tag{3}$$
$$\sigma^2 = 3 \tag{4}$$

(e)

Find the estimated variances for b_1 and b_2.

$$var(b) = \sigma^2 (X'X)^{-1} \tag{1}$$
$$= 3 \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}^{-1} \tag{2}$$
$$= 3\frac{1}{9} \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \tag{3}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4}$$

Therefore, $var(b_1) = 1$ and $var(b_2) = 1$.

(f)

Find the covariance between b_1 and b_2.

Since we found our variance-covariance matrix in (e):

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Then $cov(b_1, b_2) = 0$

(g)

What is the true variance between b_1 and b_2?

Although we found the covariance between b_1 and b_2 to be 0, we can only conclude they do not have linear dependencies. The true variance between b_1 and b_2 may be non-linearly related.

3. Exercise 7.16, page 247-248

A life insurance company wishes to examine the relationship between the amount of life insurance held by a family and family income. From a random sample of 20 households, the company collected the following observations:

Family	Life Insurance ($ thousands)	Income ($ thousands)
1	90	25
2	165	40
3	220	60
4	145	30
5	114	29
6	175	41
7	145	37
8	192	46
9	395	105
10	339	81
11	230	57
12	262	72
13	570	140
14	100	23
15	210	55
16	243	58
17	335	87
18	299	72
19	305	80
20	205	48

(a)

Estimate a linear relationship between insurance (y) and income (x).

$\hat{y} = 6.855 + 3.8802\hat{x}$

(b)

Discuss the relationship you estimated in (a). In particular:

(i)

What is your estimate of the resulting change in the amount of life insurance when income increases by $1000?

For every unit increase of income (per $1000), life insurance will increase by 3.8802 (per $1000), or by $3,880.20, ceteris paribus.

(ii)

What is the standard error of the estimate in (i) and how do you use this standard error for interval estimation and hypothesis testing?

Standard error = $\sqrt{var(b)} = \sqrt{0.0126} = 0.1122$. With the standard error, we can use it to calculate a certain confidence interval (typically around 90%, 95%, and 99%) using the formula:

$$b - t_c\sqrt{var(b)} \leq \beta_2 \leq b + t_c\sqrt{var(b)}$$

Additionally, for hypothesis testing, we can use the standard error to find our t-score, which would give us the probability that our estimator is within the true relationship:

$$t = \frac{b - \beta}{\sqrt{var(b)}}$$

(iii)

One member of the management board claims that, for every \$1000 increase in income, the amount of life insurance held will go up by \$5000. Does your estimated relationship support this claim? Use a 5% significance level.

$$H_0 : \beta_2 = 5$$
$$H_1 : \beta_2 \neq 5$$

$$t = \frac{b_2 - \beta_2}{\sqrt{var(b_2)}} \tag{1}$$
$$= \frac{3.88 - 5}{0.1122} \tag{2}$$
$$= -9.978 \tag{3}$$

For a 5% confidence interval ($\alpha = 0.05$) with degree of freedom of $20 - 2 = 18$, two-tail, $t_c = 2.1009$

Since $t < t_c \rightarrow 0.078 < 2.1009$, we fail to reject the null hypothesis.

(c)

Test the hypothesis that the amount of life insurance held is proportional to income.

$$H_0 : \beta_2 = 0$$
$$H_1 : \beta_2 \neq 0$$

$$t = \frac{b_2 - \beta_2}{\sqrt{var(b_2)}} \tag{1}$$
$$= \frac{3.88 - 0}{0.1122} \tag{2}$$
$$= 34.58 \tag{3}$$

For a 5% confidence interval ($\alpha = 0.05$) with degree of freedom of $20 - 2 = 18$, two-tail, $t_c = 2.1009$ Since $t < t_c \rightarrow 0.27 > 2.1009$, we reject the null hypothesis. This means income has some effect on the amount of life insurance held.

(d)

Predict the amount of life insurance held by a family with an income of $100,000.

$$\begin{aligned}
\hat{y} &= 6.855 + 3.8802\hat{x} && (1)\\
&= 6.855 + 3.8802(100) && (2)\\
&= 394.875 && (3)
\end{aligned}$$

The predicted amount of life insurance held by a family with an income of $100,000 is $394,875.

(e)

Ten years hence, it is found that a family with an income of $100,000 has life insurance totaling $440,000. Is there any evidence to suggest that your estimated relationship is no longer relevant?

To find the potentially new b_2, we will have to plug 440 for $\hat{y}$ and 100 for $\hat{x}$ into our previous estimated linear relationship:

$$\begin{aligned}
\hat{y} &= 6.855 + 3.8802\hat{x} && (1)\\
440 &= 6.855 + \beta_2(100) && (2)\\
\beta_2 &= \frac{440 - 6.855}{100} && (3)\\
\beta_2 &= 4.33145 && (4)
\end{aligned}$$

Now we need to test if this new β_2 is significantly different than our original:

H_0: $\beta_2 = 3.8802$
H_1: $\beta_2 \neq 3.8802$

$$\begin{aligned}
t &= \frac{b_2 - \beta_2}{\sqrt{var(b_2)}} && (1)\\
&= \frac{4.33145 - 3.88}{0.1122} && (2)\\
&= 4.0236 && (3)
\end{aligned}$$

For a 5% confidence interval ($\alpha = 0.05$) with degree of freedom of $20 - 2 = 18$, two-tail, $t_c = 2.1009$ Since $t < t_c \rightarrow 4.0236 > 2.1009$, we reject the null hypothesis. This means from this sample of a family with an income of $100,000 having a life insurance totaling $440,000 gives us evidence that our original estimator b_2 is no longer relevant.

4. Exercise 8.6, page 273-274

The catering company Thirst Quenchers has a contract to supply soda at the University of California football games. They suspect that the major factor influencing the quantity of soda consumed is the maximum temperature on the day of each game. The last three football seasons have yielded the data in Table 8.7.

Game	Soda Sold (gallons)	Max Temp, °F
1	1250	81
2	890	60
3	1093	73
4	1546	86
5	635	58
6	937	68
7	1142	75
8	1120	76
9	1067	72
10	1410	84
11	987	69
12	1198	77
13	1429	85
14	1147	74
15	1200	74
16	904	62
17	1342	83
18	1005	70

(a)

Estimate a linear equation that relates the quantity of soda sold to the maximum temperature. Construct 95% interval estimates for each of the parameters. Comment on the results.

$$\hat{y} = -771.26 + 25.76\hat{x}$$

95% confidence interval for b_1: $(-1040.77, -501.75)$

95% confidence interval for b_2: $(22.13, 29.39)$

When sampling the population, we expect 95 of our samples out of 100 to contain the true β_1 between -1040.77 and -501.75. When sampling the population, we expect 95 of our samples out of 100 to contain the true β_2 between 22.13 and 29.39.

(b)

Is there evidence to suggest that the temperature does influence the quantity consumed?

$$H_0 : \beta_2 = 0$$
$$H_1 : \beta_2 \neq 0$$

$$t = \frac{b_2 - \beta_2}{\sqrt{var(b_2)}} \tag{1}$$
$$= \frac{25.76 - 0}{\sqrt{0.0003}} \tag{2}$$
$$= 1487.25 \tag{3}$$

For $\alpha = 0.05$, at degrees of freedom: $T - K = 18 - 2 = 16$, two-tailed, then $t_c = 2.120$
Since $t > t_c \rightarrow 1487.25 > 2.120$, we reject H_0.
Yes, there is evidence to suggest that the temperature does influence the quantity of soda consumed, ceteris paribus.

(c)

At the 18 games from which the data were collected, there was always enough soda available to meet demand. Suppose that Thirst Quenchers decides to stock 110 gallons for the next game. What is the probability that they will run out of soda if the maximum temperature is (i) 70°and (ii) 75°? (To work out these probabilities assume that the coefficient vector β and the error variance σ^2 have been estimated without sampling error.)

(d)

Suppose that Thirst Quenchers can always accurately predict the maximum temperature on the day of the game. Suppose also that their decision rule is to stock 30 gallons more soda than their predicted requirements. Find the probability that they will run out of soda when the maximum temperature is (i) 70°and (ii) 80°. (Allow for sampling error in the estimation of β but assume that error variance σ^2 has been estimated without sampling error.)

1. Exercise 9.2, page 313-314

The international hamburger chain known colloquially as Makkers has recently opened a franchise in Moscow. There has been little experience with such establishments in Moscow and so there is considerable uncertainty about the optimal price for hamburgers and the optimal advertising expenditure. During the first 20 weeks of its operation, the Moscow Makkers experimented with alternative prices for its hamburgers and with the level of advertising expenditure, and collected the data on number of hamburgers sold, price, and advertising expenditure given in Table 9.5. Makkers' economist decided to model the quantity of hamburgers sold (q) as the following function of price in rubles (p) and the level of advertising expenditure in hundreds of rubles (a).

$$q_t = \beta_1 + \beta_2 p_t + \beta_3 a_t + \beta_4 a_t^2 + e_t$$

Week	Hamburgers	Prices (rubles)	Advertising (hundreds of rubles)
1	425	4.92	4.79
2	467	5.5	3.61
3	296	5.54	5.49
4	626	5.11	2.78
5	165	5.62	5.74
6	515	5.24	1.34
7	270	4.15	5.81
8	689	4.02	3.39
9	413	5.77	3.74
10	561	4.57	3.59
11	307	5.67	5.19
12	508	5.92	3.27
13	299	5.97	4.69
14	531	5.59	3.79
15	445	5.5	4.29
16	412	5.86	2.71
17	845	4.09	2.21
18	471	5.08	3.09
19	439	5.36	4.65
20	520	5.22	1.97

(a)

What quantities appear in y and X when this model is written in the matrix algebra notation $y = X\beta + e$?

$$y = \begin{bmatrix} 626 \\ 165 \\ 515 \\ 270 \\ 689 \\ 413 \\ 561 \\ 307 \\ 508 \\ 299 \\ 531 \\ 445 \\ 412 \\ 845 \\ 471 \\ 439 \\ 520 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 5.11 & 2.78 \\ 1 & 5.62 & 5.74 \\ 1 & 5.24 & 1.34 \\ 1 & 4.15 & 5.81 \\ 1 & 4.02 & 3.39 \\ 1 & 5.77 & 3.74 \\ 1 & 4.57 & 3.59 \\ 1 & 5.67 & 5.19 \\ 1 & 5.92 & 3.27 \\ 1 & 5.97 & 4.69 \\ 1 & 5.59 & 3.79 \\ 1 & 5.5 & 4.29 \\ 1 & 5.86 & 2.71 \\ 1 & 4.09 & 2.21 \\ 1 & 5.08 & 3.09 \\ 1 & 5.36 & 4.65 \\ 1 & 5.22 & 1.97 \end{bmatrix}$$

(b)

The coefficient β_2 shows the response of quantity of hamburgers sold to a change in price. It is given by the partial derivative $\partial q/\partial p$. Similarly, the response of quantity to a change in advertising expenditure is given by the partial derivative $\partial q/\partial a$. However, in this case the response is not simply equal to a constant coefficient but depends on a, the level of advertising. Specifically,

$$\frac{\partial q}{\partial p} = \beta_3 + 2\beta_4 a$$

What signs do you expect for the parameters β_2, β_3, and β_4?

We would expect β_2 to have a negative coefficient because as the price of hamburgers increases, this would cause fewer hamburgers to be sold.

On the other hand, we would expect β_3 and β_4 to have positive coefficients because as spending towards advertising increases, we expect more consumers to be aware of and want more of the hamburgers.

(c)

Find least squares for β_1, β_2, β_3, and β_4. Report these estimates and their standard errors in the conventional way. Do the signs of your estimates agree with your expectations?

$$b = \begin{bmatrix} 0.003 \\ 1.3732 \\ -0.1134 \\ -0.0839 \end{bmatrix}$$

No, the signs of my estimates do not match up with my expectations. The exact opposite happened.

(d)

Suppose that the average cost of producing hamburgers is 1 ruble and that this cost is constant (does not depend on number of hamburgers sold). Makkers' weekly profit from hamburger sales is given by

$$\text{profit} = pq - q - 100a$$

The term $100a$ arises because advertising expenditure is measured in hundreds of rubles. Use your estimated demand function to write profit as a function of p and a only.

Since:

$$\hat{q} = 0.003 + 1.3732\hat{p} - 0.1134\hat{a} - 0.0939\hat{a}^2$$

We can plug this q into the profit equation we were just given:

$$\begin{aligned}
\text{profit} &= pq - q - 100a && (1)\\
&= p(0.003 + 1.3702p - 0.1134a - 0.0939a^2) - (0.003 + 1.3732p - 0.1134a - 0.0939a^2) - 100a && (2)\\
&= -0.003 + 1.3732p^2 - 0.1134ap - 0.0839a^2p - 1.3732p + 0.0839a^2 - 99.8866a && (3)
\end{aligned}$$

(e)

Find the profit-maximizing price of hamburgers when advertising expenditure is 280 rubles. (*Hint:* Make use of the equation obtained by setting the partial derivative ∂(profit)/∂p equal to zero. The way in which partial derivations are used for obtaining maximizing values is considered in Appendix 10A.)

$$\begin{aligned}
\frac{\partial \text{profit}}{\partial p} &= \frac{\partial(0.003 + 1.3732p^2 - 0.1134ap - 0.0839a^2p - 1.3702p + 0.0839a^2 - 99.8866a)}{\partial p} && (1)\\
0 &= 2.7464p - 0.1134a - 0.0839a^2 - 1.3702 && (2)\\
0 &= 2.7464p - 0.1134(280) - 0.0839(280)^2 - 1.3702 && (3)\\
0 &= 2.7464p - 31.752 - 6,577.76 - 1.3702 && (4)\\
6,610.8822 &= 2.7464p && (5)\\
p &= 2,407.11 && (6)
\end{aligned}$$

(f)

Find the profit-maximizing level of advertising expenditure when the price of hamburgers if 5 rubles. (*Hint:* Make use of the equation obtained by setting the parital derivative ∂(profit)/∂a equal to zero.)

$$\begin{aligned}
\frac{\partial \text{profit}}{\partial a} &= \frac{\partial(0.003 + 1.3732p^2 - 0.1134ap - 0.0839a^2p - 1.3702p + 0.0839a^2 - 99.8866a)}{\partial a} && (1)\\
0 &= -0.1134p - 0.1678ap + 0.1678a - 99.8866 && (2)\\
0 &= -0.1134(5) - 0.1678a(5) + 0.1678a - 99.8866 && (3)\\
0 &= -0.567 - 0.839a + 0.1678a - 99.8866 && (4)\\
100.4536 &= -0.6712a && (5)\\
a &= -149.66 && (6)
\end{aligned}$$

(g)

Find the optimal p when $a = 2.13$. Find the optimal a when $p = 5.32$. What settings for p and a do you think Makkers' economist will recommend?

2. Exercise 10.2, page 348-349

Consider the model

$$y_t = \beta_1 + x_{t2}\beta_2 + x_{t3}\beta_3 + e_t$$

and suppose that application of least squares to 20 observations on these variables yield the following results

$$b = \begin{bmatrix} 0.96587 \\ 0.69914 \\ 1.7769 \end{bmatrix} \qquad \hat{cov}(b) = \hat{\sigma}^2(X'X)^{-1} = \begin{bmatrix} 0.21812 & 0.019195 & -0.050301 \\ 0.019195 & 0.048526 & -0.031223 \\ -0.050301 & -0.031223 & 0.037120 \end{bmatrix}$$

$$\hat{\sigma}^2 = 2.5193$$
$$R^2 = 0.9466$$

(a)

Find the total variation, unexplained variation, and explained variation for this model.

Since we know:

$$\text{Total variation = explained variation + unexplained variation}$$
$$R^2 = \frac{\text{Eplained variation}}{\text{Total variation}}$$
$$R^2 = 1 - \frac{\hat{e}'\hat{e}}{y'y - T\bar{y}^2}$$
$$\sigma^2 = \frac{\hat{e}'\hat{e}}{T-K}$$

We can attempt to find $y'y - Ty^2$ in relation to our given results:

$$\sigma^2 = \frac{\hat{e}'\hat{e}}{T-K} \quad (1)$$
$$\hat{e}'\hat{e} = \sigma^2(T-K) \quad (2)$$

$$R^2 = 1 - \frac{\hat{e}'\hat{e}}{y'y - T\bar{y}^2} \quad (1)$$
$$1 - R^2 = \frac{\hat{e}'\hat{e}}{y'y - T\bar{y}^2} \quad (2)$$
$$y'y - Ty^2 = \frac{\hat{e}'\hat{e}}{1-R^2} \quad (3)$$
$$y'y - Ty^2 = \frac{\sigma^2(T-K)}{1-R^2} \quad (4)$$

$$R^2 = 1 - \frac{\hat{e}'\hat{e}}{y'y - T\bar{y}^2} \quad (1)$$
$$R^2 = \frac{y'y - Ty^2 - \hat{e}'\hat{e}}{y'y - Ty^2} \quad (2)$$
$$R^2(y'y - Ty^2) = y'y - T\bar{y}^2 - \hat{e}'\hat{e} \quad (3)$$
$$\hat{e}'\hat{e} = y'y - Ty^2 - R^2(y'y - T\bar{y}^2) \quad (4)$$
$$(5)$$

Since from (2), we see $R^2 = \frac{y'y - Ty^2 - \hat{e}'\hat{e}}{y'y - T\bar{y}^2}$ and we know $R^2 = \frac{\text{Eplained variation}}{\text{Total variation}}$, it follows that our total variation is $y'y - Ty^2 = \frac{\sigma^2(T-K)}{1-R^2}$, our unexplained variation is $\hat{e}'\hat{e} = \sigma^2(T-K)$, and our explained variation is $R^2\frac{\sigma^2(T-K)}{1-R^2}$.

Thus,

Total variation:

$$= \frac{\sigma^2(T-K)}{1-R^2} \quad (1)$$
$$= \frac{2.5193(20-3)}{1-0.9466} \quad (2)$$
$$= \frac{42.8281}{0.0534} \quad (3)$$
$$= 802.02 \quad (4)$$

Unexplained variation:

$$
\begin{aligned}
&= \sigma^2(T-K) && (1)\\
&= 2.5193(20-3) && (2)\\
&= 42.8281 && (3)
\end{aligned}
$$

Explained variation:

$$
\begin{aligned}
&= R^2\frac{\sigma^2(T-K)}{1-R^2} && (1)\\
&= 0.9466(802.02) && (2)\\
&= 759.192132 && (3)
\end{aligned}
$$

(b)

Find 95% interval estimates for β_2 and β_3.

$$
\begin{aligned}
&b_2 - t_c\sqrt{var(b_2)} \le \beta_2 \le b_2 + t_c\sqrt{var(b_2)} && (1)\\
&0.69914 - 2.11\sqrt{0.048526} \le \beta_2 \le 0.69914 - 2.11\sqrt{0.048526} && (2)\\
&0.2343 \le \beta_2 \le 1.1639 && (3)
\end{aligned}
$$

$$
\begin{aligned}
&b_3 - t_c\sqrt{var(b_3)} \le \beta_3 \le b_3 + t_c\sqrt{var(b_3)} && (1)\\
&1.7769 - 2.11\sqrt{0.037120} \le \beta_3 \le 1.7769 - 2.11\sqrt{0.037120} && (2)\\
&1.3704 \le \beta_3 \le 2.1834 && (3)
\end{aligned}
$$

(c)

Use a t-test to test the hypothesis $H_0 : \beta_2 \ge 1$ against the alternative $H_1 : \beta_2 < 1$.

H_0: $\beta_2 \ge 1$
H_1: $\beta_2 < 1$

$$
\begin{aligned}
t &= \frac{b_2 - \beta_2}{\sqrt{var(b_2)}} && (1)\\
&= \frac{0.6914 - 1}{\sqrt{0.049526}} && (2)\\
&= -1.3867 && (3)
\end{aligned}
$$

For $\alpha = 0.05$, at degrees of freedom: $T-K = 20-3 = 17$, one-tail, then $t_c = 1.74$. Since $t < t_c \rightarrow -1.3867 < 1.74$, we fail to reject H_0.

(d)

Use your answers in part (a) to test the hypothesis that $\beta_2 = \beta_3 = 0$.

H_0: $\beta_2 = 0$
H_1: $\beta_2 \neq 0$

$$t = \frac{b_2 - \beta_2}{\sqrt{var(b_2)}} \quad (1)$$
$$= \frac{0.6914 - 0}{\sqrt{0.049526}} \quad (2)$$
$$= 3.1068 \quad (3)$$

For $\alpha = 0.05$, at degrees of freedom: $T - K = 20 - 3 = 17$, two-tail, then $t_c = 2.11$.
Since $t > t_c \rightarrow 3.1068 > 2.11$, we reject H_0 and conclude there is evidence for suggesting b_2 does have an effect in our output.

$$H_0\colon \beta_3 = 0$$
$$H_1\colon \beta_3 \neq 0$$

$$t = \frac{b_2 - \beta_2}{\sqrt{var(b_2)}} \quad (1)$$
$$= \frac{1.7769 - 0}{\sqrt{0.037120}} \quad (2)$$
$$= 9.2227 \quad (3)$$

For $\alpha = 0.05$, at degrees of freedom: $T - K = 20 - 3 = 17$, two-tail, then $t_c = 2.11$.
Since $t > t_c \rightarrow 9.2227 > 2.11$, we reject H_0 and conclude there is evidence for suggesting b_3 does have an effect in our output.

(e)

Write the hypotheses $\beta_2 = \beta_3 = 0$ in the form $\beta_s = 0$ and find $c\hat{o}v(b_s)$. Use this information to test the hypothesis. Does your answer agree with that in part (d)?

$$H_0 : \begin{cases} \beta_2 = 0 \\ \beta_3 = 0 \end{cases}$$

3.

Suppose that, from a sample of 63 observations, the least squares b and the corresponding variance-covariance matrix are give by:

$$b = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$$
$$var(b) = \begin{bmatrix} 3 & -2 & 1 \\ -2 & 4 & 0.5 \\ 1 & 0.5 & 3 \end{bmatrix}$$

Test each of the following hypotheses:

(a) $H_0 : \beta_2 = 0$

$$t = \frac{b_2 - \beta_2}{\sqrt{var(b_2)}} \quad (1)$$
$$= \frac{3 - 0}{\sqrt{4}} \quad (2)$$
$$= \frac{3}{2} \quad (3)$$

For $\alpha = 0.05$, at degrees of freedom: $T - K = 63 - 3 = 60$, two-tail, then $t_c = 2$
Since $t < t_c \rightarrow \frac{3}{2} < 2$, we fail to reject H_0.

(b) $H_0 : \beta_1 + 2\beta_2 = 5$

$$
\begin{aligned}
t &= \frac{b_1 + 2b_2 - \beta_2}{\sqrt{var(b_1 + b_2)}} && (1)\\
&= \frac{b_1 + 2b_2 - \beta_2}{\sqrt{c_1^2 var(b_1) + c_2^2 var(b_2) + 2c_1c_2cov(b_1, b_2)}} && (2)\\
&= \frac{2 + 2(3) - 5}{\sqrt{(1^2)(3) + (2^2)(4) + (2)(1)(2)(-2)}} && (3)\\
&= \frac{3}{\sqrt{3 + 16 - 8}} && (4)\\
&= \frac{3}{\sqrt{11}} && (5)\\
&= 0.904 && (6)
\end{aligned}
$$

For $\alpha = 0.05$, at degrees of freedom: $T - K = 63 - 3 = 60$, two-tail, then $t_c = 2$
Since $t < t_c \rightarrow 0.904 < 2$, we fail to reject H_0.

(c) $H_0 : \beta_1 - \beta_2 + \beta_3 = 4$

$$
\begin{aligned}
t &= \frac{b_1 - b_2 + b_3 - \beta_2}{\sqrt{var(b_1 + b_2 + b_3)}} && (1)\\
&= \frac{b_1 - b_2 + b_3 - \beta_2}{\sqrt{c_1^2 var(b_1) + c_2^2 var(b_2) + c_3^2 var(b_3) + 2c_1c_2cov(b_1, b_2) + 2c_1c_3cov(b_1, b_3) + 2c_2c_3cov(b_2, b_3)}} && (2)\\
&= \frac{2 - 3 - 1 - 4}{\sqrt{(1^2)(3) + (-1^2)(4) + (1^2)(3) + (2)(1)(-1)(-2) + (2)(1)(1)(1) + 2(-1)(1)(0.5)}} && (3)\\
&= \frac{-6}{\sqrt{3 + 4 + 3 + 4 + 2 - 1}} && (4)\\
&= \left| \frac{-6}{\sqrt{15}} \right| && (5)\\
&= 1.549 && (6)
\end{aligned}
$$

For $\alpha = 0.05$, at degrees of freedom: $T - K = 63 - 3 = 60$, two-tail, then $t_c = 2$
Since $t < t_c \rightarrow 1.549 < 2$, we fail to reject H_0.

(d) $H_0 : \begin{cases} \beta_1 + 2\beta_2 = 5 \\ \beta_1 - \beta_2 + \beta_3 = 4 \end{cases}$

Since we know:

$$F = \frac{(Rb - r)'(Rcov(b)R')^{-1}(Rb - r)}{J}$$

We can create our R and r matrices:

$$R = \begin{bmatrix} 1 & 2 & 0 \\ 1 & -1 & 1 \end{bmatrix}$$
$$r = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

Since J is the number of hypotheses we are testing, in this instance, $J = 2$

We get $F = 1.2536$.

For $\alpha = 0.05$, at degrees of freedom: $T - K = 63 - 3 = 60$, and for $J = 2$, then $F_c = 3.15$

Since $F < F_c \rightarrow 1.2536 < 3.15$, we fail to reject H_0.