AM 166 Problem Set 6 3/14/03

11.7 Note that $\sum_{i=1}^{n} x_i y_i = 134{,}542$ and $\sum_{i=1}^{n} x_i^2 = 53{,}514$. Thus

$$\widehat{\beta}_1 = \frac{134{,}542}{53{,}514} = 2.514$$

and $\widehat{y}_i = 2.514x_i$.

11.9 **a.** Using the model $y = \beta_0 + \beta_1 x + \epsilon$, calculate

$\sum_{i=1}^{14} x_i = 86.48$ $\quad\sum_{i=1}^{14} y_i = 3787$ $\quad\sum_{i=1}^{14} y_i^2 = 1{,}257{,}465$

$\sum_{i=1}^{14} x_i^2 = 732.4876$ $\quad\sum_{i=1}^{5} x_i y_i = 17{,}562.8$ $\quad S_{xy} = -5830.04$

$S_{xx} = 198.29$ $\quad\widehat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = -29.402$

$\widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1 \overline{x} = \frac{3787}{14} - (-29.402)\left(\frac{86.48}{14}\right) = 452.119$

The least squares line is $\widehat{y} = 452.119 - 29.402x$.

b. The graph is omitted.

11.11 Since $\widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1 \overline{x}$,

$$\begin{aligned}
\text{SSE} &= \sum \left[y_i - \left(\overline{y} - \widehat{\beta}_1 \overline{x}\right) - \widehat{\beta}_1 x_i\right]^2 \\
&= \sum \left[(y_i - \overline{y}) - \widehat{\beta}_1 (x_i - \overline{x})\right]^2 \\
&= \sum (y_i - \overline{y})^2 + \widehat{\beta}_1^2 \sum (x_i - \overline{x})^2 - 2\widehat{\beta}_1 \sum (x_i - \overline{x})(y_i - \overline{y}) \\
&= \sum (y_i - \overline{y})^2 + \widehat{\beta}_1 \times \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sum (x_i - \overline{x})^2} \times \sum (x_i - \overline{x})^2 \ \left(\text{as } \widehat{\beta}_1 = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sum (x_i - \overline{x})^2}\right) \\
&\quad - 2\widehat{\beta}_1 \sum (x_i - \overline{x})(y_i - \overline{y}) \\
&= \sum (y_i - \overline{y})^2 - \widehat{\beta}_1 \sum (x_i - \overline{x})(y_i - \overline{y}) = S_{yy} - \widehat{\beta}_1 S_{xy}
\end{aligned}$$

11.13a. Calculate

$$S_{yy} = \sum (y_i - \overline{y})^2 = 16{,}045.29 - \tfrac{(346.9)^2}{8} = 1002.8388$$

and

$$S_{xy} = \sum (x_i - \overline{x})(y_i - \overline{y}) = \tfrac{1626.8}{8} = 203.35.$$

Then

$$\text{SSE} = 1002.8388 - 4.84167(203.35) = 18.2858$$

and

$$s^2 = \tfrac{\text{SSE}}{6} = 3.0476.$$

b. Using the coding formula given here, we obtain the y values and the corresponding x^* values shown below.

y	27.6	32.5	35.9	39.3	44.2	48.8	55.7	62.9
x^*	−7	−5	−3	−1	1	3	5	7

Then

$\sum x^* = 0$ $\qquad \sum y = 346.9$ $\qquad \sum x^* y = 406.7$

$\sum x^{*2} = 168$ $\qquad \sum y^2 = 16{,}045.29$ $\qquad n = 8$

$S_{x^*y} = 406.7$ $\qquad S_{x^*x^*} = 168$

so that

$$\widehat{\beta}_1^* = \tfrac{S_{xy}}{S_{xx}} = 2.42$$

and

$$\widehat{\beta}_0^* = \overline{y} = 43.3625.$$

The fitted model is $\widehat{y} = 43.35 + 2.42x^*$. Calculate $\sum (y_i - \overline{y})^2 = 1002.8388$, as in part **a** , then

$$\text{SSE} = S_{xy} - \widehat{\beta}_1 S_{xx} = 1002.8388 - 2.42(406.7) = 18.2858$$

and

$$s^2 = \tfrac{\text{SSE}}{6} = 3.0476.$$

as in part **a**. Notice these are the exact same estimates we obtained in part **a** More generally, shifting the X value by a constant does not change SSE.

11.15a. Similar to previous exercises. Calculate

$\sum x_i = 160$ $\qquad \sum y_i = 106$ $\qquad \sum x_i y_i = 1848$

$\sum x_i^2 = 2880$ $\qquad \sum y_i^2 = 1236$ $\qquad n = 10$

$S_{xy} = 152.0$ $\qquad S_{xx} = 320$ $\qquad S_{yy} = 112.4$

Then

$$\widehat{\beta}_1 = \tfrac{S_{xy}}{S_{xx}} = .475$$

and

$$\widehat{\beta}_0 = \overline{y} - \widehat{\beta}\overline{x} = 10.6 - .475(1.6) = 3.000.$$

b. The least squares line, $\widehat{y} = 3.000 + 4.75x$, is shown in Figure 11.6.

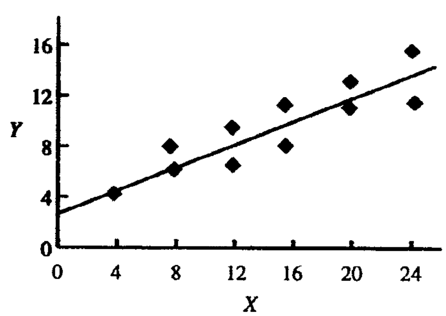

Figure 11.6

c. Calculate $S_{yy} - \widehat{\beta}_1 S_{xy} = 112.4 - .475(152) = 40.2$ and

$$s^2 = \tfrac{\text{SSE}}{n-2} = \tfrac{40.2}{8} = 5.025.$$

11.18 Refer to Exercise 11.16, in which the likelihood function is given.

$$\ln L = \ln K' - \frac{n}{2}\ln \sigma^2 - \frac{1}{2\sigma^2}\sum (y_i - \beta_0 - \beta_1 x_i)^2$$

Differentiating with respect to σ^2, we have

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum (y_i-\beta_0-\beta_1 x_i)^2}{2\sigma^4} = 0 \quad \text{or} \quad \widehat{\sigma}^2 = \frac{\sum (y_i-\beta_0-\beta_1 x_i)^2}{n}$$

Using the maximum likelihood estimates of β_0 and β_1 (see Exercise 11.16), we have

$$\widehat{\sigma}^2 = \frac{\sum (y_i-\widehat{\beta}_0-\widehat{\beta}_1 x_i)^2}{n} = \frac{\text{SSE}}{n}$$

11.20 Refer to Exercise 11.9. We wish to test

$$H_0\text{: } \beta_1 = 0 \qquad \text{vs.} \qquad H_a\text{: } \beta_1 \neq 0.$$

We have SSE $= S_{yy} - \widehat{\beta}_1 S_{xy} = 233{,}081.5 - 171{,}414.836 = 61{,}667.66$ and

$$s^2 = \frac{61{,}667.66}{12} = 5138.97.$$

The test statistic is

$$t_{12} = \frac{\widehat{\beta}_1 - 0}{s\sqrt{c_{11}}} = \frac{-29.402}{\sqrt{5138.97}\sqrt{.005043}} = -5.775$$

The p-value is $P(|t| > 5.775)$. From Table 5, $P(|t| > 3.055) = 2(.005) = .01$. Thus, the p-value $< .01$. We would reject H_0 when $\alpha = .10$.

11.21 **a.** Refer to Exercise 11.15. We wish to test

$$H_0\text{: } \beta_1 = 0 \qquad \text{vs.} \qquad H_a\text{: } \beta_1 \neq 0.$$

To carry out the test we need to estimate the value of $\sigma^2 c_{11} = V(\widehat{\beta}_1)$. From Exercise 11.15, $s^2 = \widehat{\sigma}^2 = 5.025$. From Section 11.4, $V(\widehat{\beta}_1) = \frac{\sigma^2}{\sum (x_i - \overline{x})^2}$. Now

$$\sum (x_i - \overline{x}) = 2880 - \frac{(160)^2}{10} = 320 = \frac{1}{c_{11}}$$

Thus,

$$V(\widehat{\beta}_1) = \frac{\sigma^2}{320} = \frac{5.025}{30}$$

The test statistic is

$$t = \frac{\widehat{\beta}_1}{\sqrt{\frac{s^2}{\sum (x_i-\overline{x})^2}}} = \frac{.475}{\sqrt{\frac{5.025}{320}}} = 3.791$$

The p-value is $P(|t_8| > 3.791) < P(|t_8| > 3.355) = 2(.005) = .01$ (see Table 5).

b. We would reject H_0 when $\alpha = .05$, since the p-value $< .05$.

c. No, there is no reason to believe that the linear trend will continue beyond the region of data collected. For example, we might would expect the rate of increase in the number of mistakes to level off at some point.

d. A 95% confidence interval for β_i is

$$\widehat{\beta}_1 \pm t_{.025}\, s\sqrt{c_{11}} \quad \text{or} \quad .475 \pm 2.306\sqrt{\frac{5.025}{320}} \quad \text{or} \quad .475 \pm .289$$

We are 95% confident that the change in the number of errors per hour of sleep depravation lies in this interval. Further, he lower endpoint being larger than 0 suggest a statistically significant positive linear increase.

11.22a. $\sum_{i=1}^{6} x_i = 323.4$ $\sum_{i=1}^{6} y_i = 42.6$ $\sum_{i=1}^{6} x_i y_i = 2495.08$

$\sum_{i=1}^{6} x_i^2 = 19{,}111.95$ $\sum_{i=1}^{6} y_i^2 = 326.06$ $S_{xy} = 198.94$

$S_{xx} = 1680.69$ $S_{yy} = 23.6$

$\widehat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{198.94}{1680.69} = .118.$

$\widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1 \overline{x} = \frac{42.6}{6} - .118\left(\frac{323.4}{6}\right) = .72.$

b. $\text{SSE} = S_{yy} - \widehat{\beta}_1 S_{xy} = 23.6 - (.118)(198.94) = .125$

and

$s^2 = \frac{\text{SSE}}{n-2} = \frac{.052}{6-2} = .013$

A 95% confidence interval for β_1 is

$$\widehat{\beta}_1 \pm t_{.025,4}\, s\sqrt{c_{11}}$$
$$.118 \pm 2.776\sqrt{.013}\ \sqrt{.00059}$$
$$.118 \pm .008$$

c. When $x = 0$, $E(Y) = \beta_0 + \beta_1(0) = \beta_0$. Thus, we must test

$H_0\colon\ \beta_0 = 0$ vs. $H_a\colon\ \beta_0 \neq 0.$

The test statistic is

$$t = \frac{\widehat{\beta}_0}{s\sqrt{c_{00}}} = \frac{.72}{\sqrt{.013}\ \sqrt{1.895}} = 4.587$$

From Table 5, $3.747 < 4.587 < 4.604$ so that the p-value is between $2(.01)$ and $2(.005)$. Since the p-value $< .05$, we would reject H_0 at the $\alpha = .05$ level.