Applied Math 166 Problem Set 7 4/2/03

11.24 Restricting ourselves to Ω_0, we find the likelihood function to be

$$L(\Omega_0) = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0)^2\right]$$

One may verify that maximum likelihood estimates of β_0 and σ^2 are

$\widehat{\beta}_0 = \overline{Y}$ and $\widehat{\sigma}_0^2 = \frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{n}$ so that

$$L(\widehat{\Omega}_0) = \frac{1}{(2\pi)^{n/2}\left(\widehat{\sigma}_0^2\right)^{n/2}} e^{-n/2}.$$

For β_1 in the unrestricted space Ω, the likelihood function is given in the solution to Exercise 11.16, and the maximum likelihood estimates of β_0 and β_1 are the least squares estimates

$$\widehat{\beta}_0 = \overline{Y} - \widehat{\beta}_1\overline{x} \qquad \text{and} \qquad \widehat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$$

The maximum likelihood estimate of σ^2 is $\widehat{\sigma}^2 = \frac{\sum_{i=1}^{n}(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i)^2}{n}$ so that

$$L(\widehat{\Omega}) = \frac{1}{(2\pi)^{n/2}\left(\widehat{\sigma}^2\right)^{n/2}} e^{-n/2}.$$

Hence

$$\lambda^{2/n} = \frac{\widehat{\sigma}^2}{\widehat{\sigma}_0^2} = \frac{\sum_{i=1}^{n}(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i)^2}{S_{yy}} = \frac{\text{SSE}}{S_{yy}}$$

Then

$$\begin{aligned} S_{yy} &= \sum_{i=1}^{n}\left[y_i - \overline{y} - \widehat{\beta}_1(x_i - \overline{x}) + \widehat{\beta}_1(x_i - \overline{x})^2\right] \\ &= \sum_{i=1}^{n}\left[y_i - \overline{y} - \widehat{\beta}_1(x_i - \overline{x})^2\right] + \widehat{\beta}_1^2 S_{xx} + 2\widehat{\beta}_1 S_{xy} - 2\widehat{\beta}_1^2 S_{xx} \\ &= \text{SSE} + 2\widehat{\beta}_1 S_{xy} - \widehat{\beta}_1^2 S_{xx} \end{aligned}$$

But $\widehat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$ so that

$$S_{yy} = \text{SSE} + 2\,\frac{S_{xy}^2}{S_{xx}} - \frac{S_{xy}^2}{S_{xx}} = \text{SSE} + \frac{S_{xy}^2}{S_{xx}}.$$

Now

$$\lambda^{2/n} = \frac{\text{SSE}}{\text{SSE} + \left(\frac{S_{xy}^2}{S_{xx}}\right)} = \frac{1}{1 + \left(\frac{T^2}{n-2}\right)}$$

where $T^2 = \frac{S_{xy}^2}{S_{xx}\left(\frac{\text{SSE}}{n-2}\right)} = \frac{\beta_1^2 \sum_i (x_i - \overline{x})^2}{\frac{\text{SSE}}{n-2}}$. Hence as λ gets small, T^2 will get large (either positively or negatively), and rejecting H_0 for large or small values of T will be equivalent to rejecting H_0 for small values of λ. Note that

$$T = \frac{\widehat{\beta}_1}{\sqrt{\frac{\widehat{\sigma}^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}}}$$

is the t test given in Section 11.6.

11.27 Using the coding $x = \frac{\text{year}-1971.5}{.5}$, we obtain the following calculations:

$\sum x = 0$ $\quad \sum y = 215.9$ $\quad \sum xy = -174.9$

$\sum x^2 = 330$ $\quad \sum y^2 = 4760.43$ $\quad n = 10$

$S_{xy} = -174.9$ $\quad S_{xx} = 330$ $\quad S_{yy} = 99.149$

Then $\widehat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{-174.9}{330} = -.53$ and

$$\text{SSE} = 99.149 - (-.53)(-174.9) = 6.452$$

and

$$s^2 = \frac{\text{SSE}}{8} = .8065.$$

To test the hypothesis

$$H_0\colon \beta_1 = 0 \qquad \text{vs.} \qquad H_a\colon \beta_1 < 0$$

we use the test statistic

$$t = \frac{\widehat{\beta}_1}{\sqrt{\frac{s^2}{\sum(x_i-\overline{x})^2}}} = \frac{-.53}{\sqrt{\frac{.8065}{330}}} = -10.72.$$

The rejection region with $\alpha = .05$ is $t < -1.86$ and H_0 is rejected. We conclude that the rate of tuberculosis is decreasing with time.

11.28 **a.** Using the uncoded x's given in Exercise 11.3, we obtain $\widehat{\beta}_1 = 4.8417$ and $S_{xx} = 42$. Then from Exercise 11.13, we obtain $s^2 = 3.0476$. To test the hypothesis

$$H_0\colon \beta_1 = 0 \qquad \text{vs.} \qquad H_a\colon \beta_1 > 0$$

we use the test statistic

$$t = \frac{\widehat{\beta}_1 - 0}{\sqrt{\frac{s^2}{S_{xx}}}} = \frac{4.841667}{\sqrt{\frac{3.0476}{42}}} = 17.97$$

The rejection region is $t > t_{.01,6} = 3.143$, and H_0 is rejected. There is evidence of an increase.

b. The 99% confidence interval for β_1 is

$$\widehat{\beta}_1 \pm t_{.005}\sqrt{\frac{s^2}{S_{xx}}} = 4.84 \pm 3.707(.26937) = 4.84 \pm 1.00,$$

or $[3.84, 5.84]$.

11.31 $V\left(a_0\widehat{\beta}_0 + a_1\widehat{\beta}_1\right) = \left(\frac{a_0^2\frac{\sum x_i^2}{n} + a_1^2 - 2a_0a_1\overline{x}}{S_{xx}}\right)$

Letting $a_0 = 1$ and $a_1 = x^*$, we have

$$V\left(\widehat{\beta}_0 + \widehat{\beta}_1 x^*\right) = \frac{\frac{\sum x_i^2}{n} + (x^*)^2 - 2x^*\overline{x}}{S_{xx}} = \frac{\frac{\sum x_i^2 - \frac{1}{n}(\sum x_i)^2}{n} + (x^*)^2 - 2x^*\overline{x} + \overline{x}^2}{S_{xy}}$$

$$= \frac{\frac{S_{xx}}{n} + (x^* - \overline{x})^2}{S_{xx}} = \frac{1}{n} + \frac{(x^* - \overline{x})^2}{S_{xx}}.$$

Since $(x^* - \overline{x})^2 \geq 0$ for all x^*, $V\left(\widehat{\beta}_0 + \widehat{\beta}_1 x^*\right)$ is minimized for $(x^* - \overline{x})^2 = 0$ or $x^* = \overline{x}$.

11.32 Refer to Exercises 11.13, 11.20, and 11.31. When $x^* = 5$, $\widehat{y} = 452.119 - 29.402(5) = 305.11$. From Exercise 11.31,

$$V(\widehat{y}) = V\left(\widehat{\beta}_0 + \widehat{\beta}_1 x\right) = \left[\frac{1}{14} + \frac{\left(5 - \frac{86.48}{14}\right)^2}{732.4876 - \left(\frac{(86.48)^2}{14}\right)}\right] s^2 = \left(\frac{1}{14} + \frac{1.3857}{198.288}\right)(5138.97)$$

$$= 402.98$$

Thus, a 90% confidence interval for $E(Y)$ is

$$\widehat{y} \pm t_{.05,12}\sqrt{V(\widehat{y})} = 305.11 \pm 1.782\sqrt{402.98} = 305.11 \pm 35.773.$$

11.38 In Exercise 11.31, we showed that

$$V(\widehat{\beta}_0 + \widehat{\beta}_1 x) = \frac{1}{n} + \frac{(x-\overline{x})^2}{S_{xx}}.$$

Hence, the prediction interval can be written as

$$\widehat{y} \pm t_{\alpha/2}\, s\sqrt{1 + \frac{1}{n} + \frac{(x-\overline{x})^2}{S_{xx}}}.$$

as shown in the text. When $x = \overline{x}$, the length of this interval is

$$2t_{\alpha/2}\, s\sqrt{1 + \frac{1}{n}}.$$

which is the minimum value the length can take on. Notice as long as the confidence coefficient stays the same, the width of the interval is controlled by $\frac{(x-\overline{x})^2}{S_{xx}}$, which is minimized at $x = \overline{x}$. I.e. our prediction will be the best at the mean of the independent variable.

11.39 Refer to Exercises 11.3 and 11.13. When $x = 9, \widehat{y} = 65.15$, and the 95% prediction interval is

$$65.15 \pm 2.447\sqrt{3.05\left[1 + \frac{1}{8} + \frac{(9-4.5)^2}{42}\right]} = 65.15 \pm 5.42$$

or $[59.73, 70.57]$.