Stat 252.01 Winter 2005
Assignment #8 Solutions

(11.17) Using standard properties of covariances, and the formulæ derived in class for $\hat{\beta}_0$, $\hat{\beta}_1$, we compute

$$\begin{aligned}
\operatorname{Cov}(\hat{\beta}_0, \hat{\beta}_1) &= \operatorname{Cov}(\overline{Y} - \hat{\beta}_1 \overline{x}, \hat{\beta}_1) \\
&= \operatorname{Cov}(\overline{Y}, \hat{\beta}_1) - \overline{x} \operatorname{Cov}(\hat{\beta}_1, \hat{\beta}_1) \\
&= \operatorname{Cov}\left(\frac{1}{n}\sum_i Y_i, \frac{1}{S_{xx}}\sum_i (x_i - \overline{x})Y_i\right) - \overline{x}\operatorname{Var}(\hat{\beta}_1) \\
&= \frac{1}{nS_{xx}} \operatorname{Cov}\left(\sum_i Y_i, \sum_i (x_i - \overline{x})Y_i\right) - \overline{x}\operatorname{Var}(\hat{\beta}_1) \\
&= \frac{1}{nS_{xx}} \sum_i \sum_j (x_j - \overline{x}) \operatorname{Cov}(Y_i, Y_j) - \overline{x}\operatorname{Var}(\hat{\beta}_1) \\
&= \frac{1}{nS_{xx}} \left[\sum_i (x_i - \overline{x}) \operatorname{Var}(Y_i) + \sum_{i \neq j} (x_j - \overline{x}) \operatorname{Cov}(Y_i, Y_j)\right] - \overline{x}\operatorname{Var}(\hat{\beta}_1)
\end{aligned}$$

Since $\operatorname{Var}(Y_i) = \sigma^2$, we conclude that

$$\sum_i (x_i - \overline{x}) \operatorname{Var}(Y_i) = \sigma^2 \sum_i (x_i - \overline{x}) = 0.$$

Since Y_i and Y_j are independent for $i \neq j$, we conclude $\operatorname{Cov}(Y_i, Y_j) = 0$ for $i \neq j$. We found in class that

$$\operatorname{Var}(\hat{\beta}_1) = \frac{1}{S_{xx}}\sigma^2.$$

Together these imply that

$$\operatorname{Cov}(\hat{\beta}_0, \hat{\beta}_1) = \frac{-\overline{x}}{S_{xx}}\sigma^2.$$

If $\sum_i x_i = 0$, then clearly $\operatorname{Cov}(\hat{\beta}_0, \hat{\beta}_1) = 0$. It follows that $\hat{\beta}_0$ and $\hat{\beta}_1$ will be independent provided that both $\hat{\beta}_0$ and $\hat{\beta}_1$ are normally distributed. Recall that linear combinations of independent normal random variables are normal. That is, if $A_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ and the A_i are independent, then $A_1 + \cdots + A_n \sim \mathcal{N}(\mu_1 + \cdots + \mu_n, \sigma_1^2 + \cdots + \sigma_n^2)$. Since $Y_i \sim \mathcal{N}(\beta_0 + \beta_1 x_i, \sigma^2)$, we see that

$$\hat{\beta}_1 = \frac{1}{S_{xx}}\sum_i (x_i - \overline{x})Y_i \sim \mathcal{N}\left(\frac{1}{S_{xx}}\sum_i (x_i - \overline{x})(\beta_0 + \beta_1 x_i), \frac{\sigma^2}{S_{xx}^2}\sum_i (x_i - \overline{x})^2\right) = \mathcal{N}\left(\beta_1, \frac{\sigma^2}{S_{xx}}\right).$$

and

$$\begin{aligned}
\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{x} &\sim \mathcal{N}\left(\beta_0 + \beta_1 \overline{x} - \frac{\overline{x}}{S_{xx}}\sum_i (x_i - \overline{x})(\beta_0 + \beta_1 x_i), \frac{\sigma^2}{n} + \frac{\sigma^2}{S_{xx}^2}\sum_i (x_i - \overline{x})^2\right) \\
&= \mathcal{N}\left(\beta_0, \frac{\sum x_i^2}{nS_{xx}}\right).
\end{aligned}$$

The independence now follows. (*Look again at the box on page 550. Check out* $\mathbb{E}(\hat{\beta}_i)$ *and* $\operatorname{Var}(\hat{\beta}_i)$. *Funny that!*)

(11.22) **(a)** We compute

$$\sum_{i=1}^{6} x_i = 323.4, \quad \sum_{i=1}^{6} x_i^2 = 19\,111.96, \quad \sum_{i=1}^{6} y_i = 42.6, \quad \sum_{i=1}^{6} y_i^2 = 326.06, \quad \sum_{i=1}^{6} x_i y_i = 2495.08,$$

and

$$S_{xy} = 198.94, \quad S_{xx} = 1680.7, \quad S_{yy} = 23.6,$$

so that

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{198.94}{1680.7} = \frac{29}{245} \approx 0.118,$$

and

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x} = \frac{42.6}{6} - \frac{29}{245} \cdot \frac{323.4}{6} = 0.72.$$

In other words, the least squares line is given by

$$\hat{y} = \frac{18}{25} + \frac{29}{245}\,x.$$

(b) We compute

$$\text{SSE} = S_{yy} - \hat{\beta}_1 S_{xy} = 23.6 - 0.118 \cdot 198.94 = 0.052,$$

and

$$s^2 = \frac{\text{SSE}}{n-2} = \frac{0.052}{6-2} = 0.013.$$

Therefore, an approximate 95% CI for $\hat{\beta}_1$ is

$$\hat{\beta}_1 \pm t_{4,0.025}\, s \sqrt{c_{11}}$$

or

$$0.118 \pm 2.776 \cdot \sqrt{0.013} \cdot \sqrt{0.000595}$$

or, approximately,

$$0.118 \pm 0.008.$$

(c) When $x = 0$, we find $\mathbb{E}(Y) = \beta_0$. Hence, we must test $H_0 : \beta_0 = 0$ against $H_A : \beta_0 \neq 0$. The test statistic is given by

$$t = \frac{\hat{\beta}_1}{s\sqrt{c_{00}}} \approx 4.587.$$

From Table 5, we find that $3.747 < 4.587 < 4.606$ so that the p-value is bracketed between $2 \cdot 0.01$ and $2 \cdot 0.005$. (It is a two-sided alternative.) Since the p-value is necessarily smaller than 0.05, we reject H_0 at the $\alpha = 0.05$ level.

(11.23) **(a)** Assuming that $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, we find from Exercise 11.17 that the distribution of $\hat{\beta}_i$ for $i = 0$, 1, is

$$\hat{\beta}_i \sim \mathcal{N}(\beta_i, c_{ii}\sigma^2).$$

Hence, under H_0,

$$Z = \frac{\hat{\beta}_i - \beta_{i0}}{\sqrt{c_{ii}}\,\sigma} \sim \mathcal{N}(0, 1).$$

As noted on page 550,

$$W = \frac{(n-2)S^2}{\sigma^2} \sim \chi^2_{n-2}.$$

Thus, by Definition 7.2,

$$T = \frac{Z}{\sqrt{W/(n-2)}} = \frac{\frac{\hat{\beta}_i - \beta_{i0}}{\sqrt{c_{ii}}\,\sigma}}{\sqrt{\frac{(n-2)S^2}{\sigma^2}/(n-2)}} = \frac{\hat{\beta}_i - \beta_{i0}}{S\sqrt{c_{ii}}}$$

has a t distribution with $n-2$ degrees of freedom.

(b) The interpretation of a $(1-\alpha)$ confidence interval is that

$$\alpha = P\left(-t_{\alpha/2,n-2} \le T \le t_{\alpha/2,n-2}\right).$$

Hence, from part (a),

$$\alpha = P\left(-t_{\alpha/2,n-2} \le \frac{\hat{\beta}_i - \beta_i}{S\sqrt{c_{ii}}} \le t_{\alpha/2,n-2}\right)$$

so that

$$\alpha = P\left(\hat{\beta}_i - t_{\alpha/2,n-2}S\sqrt{c_{ii}} \le \beta_i \le \hat{\beta}_i + t_{\alpha/2,n-2}S\sqrt{c_{ii}}\right).$$

In other words, a $(1-\alpha)$ confidence interval for β_i is given by

$$\hat{\beta}_i \pm t_{\alpha/2,n-2}S\sqrt{c_{ii}}.$$

11.31 Using our results in (11.17), we find

$$\begin{aligned}
\mathrm{Var}(\hat{\beta}_0 + \hat{\beta}_1 x^*) &= \mathrm{Var}(\hat{\beta}_0) + (x^*)^2\,\mathrm{Var}(\hat{\beta}_1) + 2x^*\,\mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1) = \frac{\sum x_i^2}{nS_{xx}}\sigma^2 + (x^*)^2\frac{\sigma^2}{S_{xx}} - 2x^*\frac{\overline{x}}{S_{xx}}\sigma^2 \\
&= \frac{\sigma^2}{S_{xx}}\left[\frac{\sum x_i^2}{n} + (x^*)^2 - 2x^*\overline{x}\right] = \frac{\sigma^2}{S_{xx}}\left[\frac{\sum x_i^2}{n} + (x^*)^2 - 2x^*\overline{x} + \overline{x}^2 - \overline{x}^2\right] \\
&= \frac{\sigma^2}{S_{xx}}\left[\frac{\sum x_i^2}{n} - \overline{x}^2 + (x^* - \overline{x})^2\right] = \frac{\sigma^2}{S_{xx}}\left[\frac{\sum x_i^2 - n\overline{x}^2}{n} + (x^* - \overline{x})^2\right] \\
&= \frac{\sigma^2}{S_{xx}}\left[\frac{S_{xx}}{n} + (x^* - \overline{x})^2\right] = \left[\frac{1}{n} + \frac{(x^* - \overline{x})^2}{S_{xx}}\right]\sigma^2
\end{aligned}$$

as required. The confidence interval for $\mathbb{E}(Y)$ achieves its shortest length when the variance of $\hat{\beta}_0 + \hat{\beta}_1 x^*$ is as small as possible. (WHY?) From the formula we just derived, this obviously occurs when $x^* = \overline{x}$.