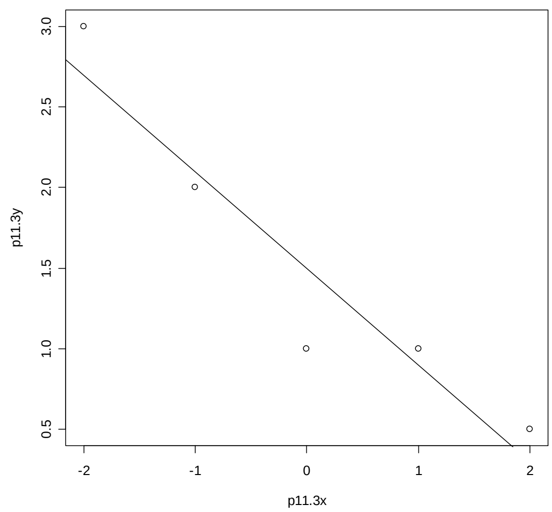
Chapter 11: Linear Models and Estimation by Least Squares

11.1 Using the hint, $\hat{y}(\bar{x}) = \hat{\beta}_0 + \hat{\beta}_1\bar{x} = (\bar{y} - \hat{\beta}_1\bar{x}) + \hat{\beta}_1\bar{x} = \bar{y}$.

11.2 **a.** slope = 0, intercept = 1. SSE = 6.
b. The line with a negative slope should exhibit a better fit.
c. SSE decreases when the slope changes from .8 to .7. The line is pivoting around the point (0, 1), and this is consistent with ($\bar{x}$, $\bar{y}$) from part Ex. 11.1.
d. The best fit is: $y = 1.000 + 0.700x$.

11.3 The summary statistics are: $\bar{x} = 0$, $\bar{y} = 1.5$, $S_{xy} = -6$, $S_{xx} = 10$. Thus, $\hat{y} = 1.5 - .6x$.



The graph is above.

11.4 The summary statistics are: $\bar{x} = 72$, $\bar{y} = 72.1$, $S_{xy} = 54{,}243$, $S_{xx} = 54{,}714$. Thus, $\hat{y} = 0.72 + 0.99x$. When $x = 100$, the best estimate of y is $\hat{y} = 0.72 + 0.99(100) = 99.72$.

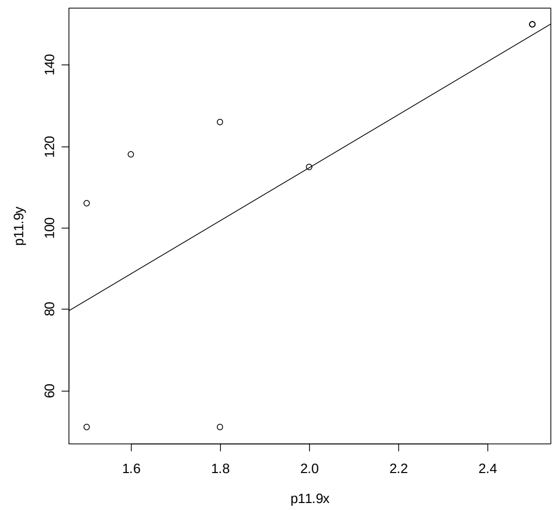
11.5 The summary statistics are: $\bar{x} = 4.5$, $\bar{y} = 43.3625$, $S_{xy} = 203.35$, $S_{xx} = 42$. Thus, $\hat{y} = 21.575 + 4.842x$. Since the slope is positive, this suggests an increase in median prices over time. Also, the expected annual increase is \$4,842.

11.6 **a.** intercept = 43.362, SSE = 1002.839.
b. the data show an increasing trend, so a line with a negative slope would not fit well.
c. Answers vary.
d. Answers vary.
e. (4.5, 43.3625)
f. The sum of the areas is the SSE.

11.7 **a.** The relationship appears to be proportional to x^2.
b. No.
c. No, it is the best *linear* model.

11.8 The summary statistics are: $\bar{x} = 15.505$, $\bar{y} = 9.448$, $S_{xy} = 1546.459$, $S_{xx} = 2359.929$. Thus, $\hat{y} = -0.712 + 0.655x$. When $x = 12$, the best estimate of y is $\hat{y} = -.712 + 0.655(12) = 7.148$.

11.9 **a.** See part **c**.
b. $\hat{y} = -15.45 + 65.17x$.

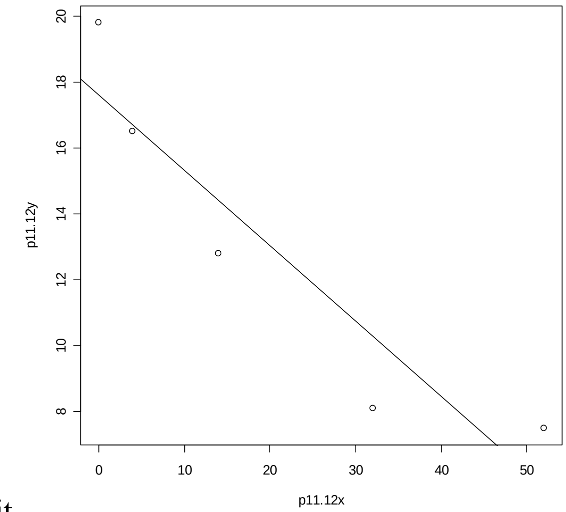


c. The graph is above.
d. When $x = 1.9$, the best estimate of y is $\hat{y} = -15.45 + 65.17(1.9) = 108.373$.

11.10 $\dfrac{d\text{SSE}}{d\beta_1} = -2\sum_{i=1}^{n}(y_i - \beta_1 x_i)x_i = -2\sum_{i=1}^{n}(x_i y_i - \beta_1 x_i^2) = 0$, so $\hat{\beta}_1 = \dfrac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$.

11.11 Since $\sum_{i=1}^{n} x_i y_i = 134{,}542$ and $\sum_{i=1}^{n} x_i^2 = 53{,}514$, $\hat{\beta}_1 = 2.514$.

11.12 The summary statistics are: $\bar{x} = 20.4$, $\bar{y} = 12.94$, $S_{xy} = -425.571$, $S_{xx} = 1859.2$.
a. The least squares line is: $\hat{y} = 17.609 - 0.229x$.

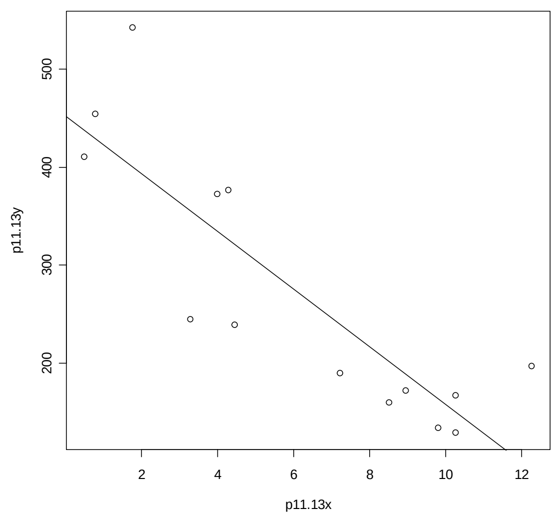


b. The line provides a reasonable fit.

c. When $x = 20$, the best estimate of y is $\hat{y} = 17.609 - 0.229(20) = 13.029$ lbs.

11.13 The summary statistics are: $\bar{x} = 6.177$, $\bar{y} = 270.5$, $S_{xy} = -5830.04$, $S_{xx} = 198.29$.

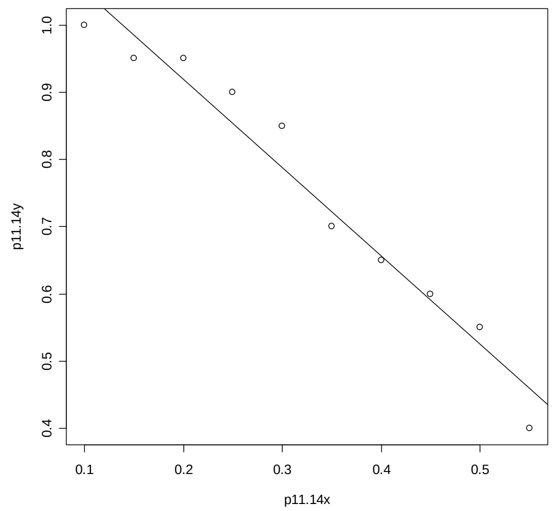
a. The least squares line is: $\hat{y} = 452.119 - 29.402x$.



b. The graph is above.

11.14 The summary statistics are: $\bar{x} = .325$, $\bar{y} = .755$, $S_{xy} = -.27125$, $S_{xx} = .20625$

a. The least squares line is: $\hat{y} = 1.182 - 1.315x$.



b. The graph is above. The line provides a reasonable fit to the data.

11.15 **a.** SSE $= \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 = \sum_{i=1}^{n}[y_i - \bar{y} - \hat{\beta}_1(x_i - \bar{x})]^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2)$

$$+ \hat{\beta}_1^2 \sum_{i=1}^{n}(x_i - \bar{x})^2 - 2\hat{\beta}_1 \sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})$$

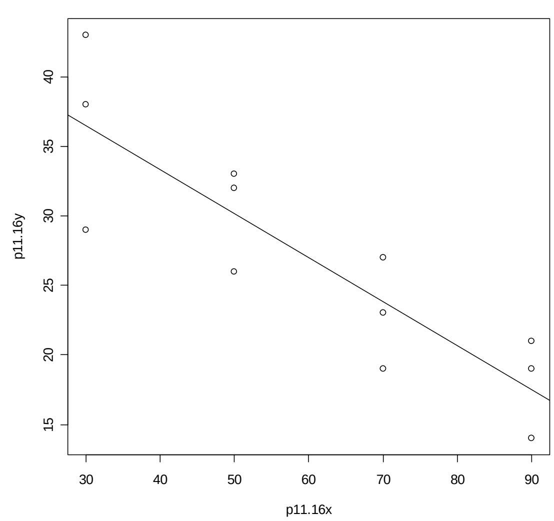
$$= \sum_{i=1}^{n}(y_i - \bar{y})^2 + \hat{\beta}_1 S_{xy} - 2\hat{\beta}_1 S_{xy} = S_{yy} - \hat{\beta}_1 S_{xy}.$$

b. Since $\text{SSE} = S_{yy} - \hat{\beta}_1 S_{xy}$,

$S_{yy} = \text{SSE} + \hat{\beta}_1 S_{xy} = \text{SSE} + (S_{xy})^2 / S_{xx}$. But, $S_{xx} > 0$ and $(S_{xy})^2 \geq 0$. So, $S_{yy} \geq \text{SSE}$.

11.16 The summary statistics are: $\bar{x} = 60$, $\bar{y} = 27$, $S_{xy} = -1900$, $S_{xx} = 6000$.

a. The least squares line is: $\hat{y} = 46.0 - .31667x$.



b. The graph is above.

c. Using the result in Ex. 11.15(a), SSE = $S_{yy} - \hat{\beta}_1 S_{xy}$ = 792 – (–.31667)(–1900) = 190.327. So, s^2 = 190.327/10 = 19.033.

11.17 **a.** With S_{yy} = 1002.8388 and S_{xy} = 203.35, SSE = 1002.8388 – 4.842(203.35) = 18.286. So, s^2 = 18.286/6 = 3.048.

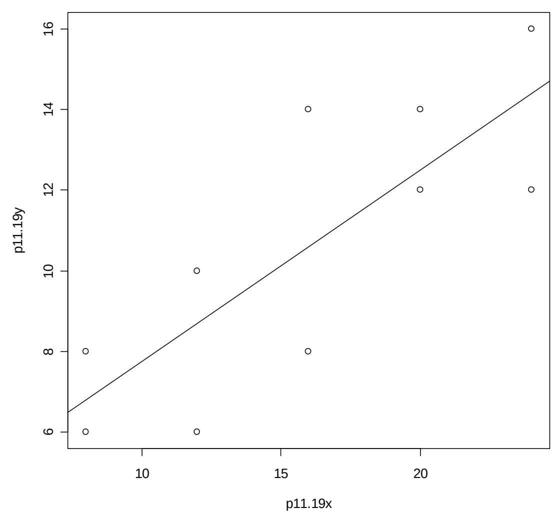
b. The fitted line is $\hat{y}$ = 43.35 + 2.42x^*. The same answer for SSE (and thus s^2) is found.

11.18 **a.** For Ex. 11.8, S_{yy} = 1101.1686 and S_{xy} = 1546.459, SSE = 1101.1686 – .6552528(1546.459) = 87.84701. So, s^2 = 87.84701/8 = 10.98.

b. Using the coding $x_i^* = x_i - \bar{x}$, the fitted line is $\hat{y}$ = 9.448 + .655x^*. The same answer for s^2 is found.

11.19 The summary statistics are: $\bar{x}$ = 16, $\bar{y}$ = 10.6, S_{xy} = 152.0, S_{xx} = 320.

a. The least squares line is: $\hat{y}$ = 3.00 + 4.75x.



b. The graph is above.

c. s^2 = 5.025.

11.20 The likelihood function is given by, $K = \left(\sigma\sqrt{2\pi}\right)^n$,

$$L(\beta_0, \beta_1) = K \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2\right], \text{ so that}$$

$$\ln L(\beta_0, \beta_1) = \ln K - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 .$$

Note that maximizing the likelihood (or equivalently the log–likelihood) with respect to β_0 and β_1 is identical to minimizing the positive quantity $\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$. This is the least–squares criterion, so the estimators will be the same.

11.21 Using the results of this section and Theorem 5.12,

$$\text{Cov}(\hat\beta_0,\hat\beta_1) = \text{Cov}(\overline{Y} - \hat\beta_1\overline{x},\hat\beta_1) = \text{Cov}(\overline{Y},\hat\beta_1) - \text{Cov}(\hat\beta_1\overline{x},\hat\beta_1) = 0 - \overline{x}V(\hat\beta_1).$$

Thus, $\text{Cov}(\hat\beta_0,\hat\beta_1) = -\overline{x}\sigma^2/S_{xx}$. Note that if $\sum_{i=1}^n x_i = 0$, $\overline{x} = 0$ so $\text{Cov}(\hat\beta_0,\hat\beta_1) = 0$.

11.22 From Ex. 11.20, let $\theta = \sigma^2$ so that the log–likelihood is

$$\ln L(\theta) = -\tfrac{n}{2}\ln(2\pi) - \tfrac{n}{2}\ln\theta - \frac{1}{2\theta}\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 .$$

Thus,

$$\tfrac{d}{d\theta}\ln L(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 .$$

The MLE is $\hat\theta = \hat\sigma^2 = \tfrac{1}{n}\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$, but since β_0 and β_1 are unknown, we can insert their MLEs from Ex. 11.20 to obtain:

$$\hat\sigma^2 = \tfrac{1}{n}\sum_{i=1}^n \left(y_i - \hat\beta_0 - \hat\beta_1 x_i\right)^2 = \tfrac{1}{n}\text{SSE}.$$

11.23 From Ex. 11.3, it is found that $S_{yy} = 4.0$

a. Since SSE = 4 – (–.6)(–6) = .4, $s^2 = .4/3 = .1333$. To test H_0: $\beta_1 = 0$ vs. H_a: $\beta_1 \neq 0$, $|t| = \frac{|-.6|}{\sqrt{.1333(.1)}} = 5.20$ with 3 degrees of freedom. Since $t_{.025} = 3.182$, we can reject H_0.

b. Since $t_{.005} = 5.841$ and $t_{.01} = 4.541$, $.01 < p$–value $< .02$.
Using the Applet, $2P(T > 5.20) = 2(.00691) = .01382$.

c. $-.6 \pm 3.182\sqrt{.1333}\sqrt{.1} = -.6 \pm .367$ or (–.967, –.233).

11.24 To test H_0: $\beta_1 = 0$ vs. H_a: $\beta_1 \neq 0$, SSE = 61,667.66 and $s^2 = 5138.97$. Then, $|t| = \frac{|-29.402|}{\sqrt{5139.97(.005043)}} = 5.775$ with 12 degrees of freedom.

a. From Table 5, $P(|T| > 3.055) = 2(.005) = .01 > p$–value.

b. Using the Applet, $2P(T > 5.775) = .00008$.

c. Reject H_0.

11.25 From Ex. 11.19, to test H_0: $\beta_1 = 0$ vs. H_a: $\beta_1 \neq 0$, $s^2 = 5.025$ and $S_{xx} = 320$. Then, $|t| = \frac{|.475|}{\sqrt{5.025/320}} = 3.791$ with 8 degrees of freedom.

a. From Table 5, $P(|T| > 3.355) = 2(.005) = .01 > p$–value.

b. Using the Applet, $2P(T > 3.791) = 2(.00265) = .0053$.

c. Reject H_0.

d. We cannot assume the linear trend continues – the number of errors could level off at some point.

e. A 95% CI for β_1: $.475 \pm 2.306\sqrt{5.025/320} = .475 \pm .289$ or (.186, .764). We are 95% confident that the expected change in number of errors for an hour increase of lost sleep is between (.186, .764).

11.26 The summary statistics are: $\overline{x} = 53.9$, $\overline{y} = 7.1$, $S_{xy} = 198.94$, $S_{xx} = 1680.69$, $S_{yy} = 23.6$.

a. The least squares line is: $\hat{y} = 0.72 + 0.118x$.

b. SSE = 23.6 – .118(198.94) = .125 so s^2 = .013. A 95% CI for β_1 is 0.118 ± $2.776\sqrt{.013}\sqrt{.00059}$ = 0.118 ± .008.

c. When x = 0, $E(Y) = \beta_0 + \beta_1(0) = \beta_0$. So, to test H_0: $\beta_0 = 0$ vs. H_a: $\beta_0 \neq 0$, the test statistic is $|t| = \frac{|.72|}{\sqrt{.013}\sqrt{1.895}}$ = 4.587 with 4 degrees of freedom. Since $t_{.005}$ = 4.604 and $t_{.01}$ = 3.747, we know that .01 < p–value < .02.

d. Using the Applet, $2P(T > 4.587) = 2(.00506) = .01012$.

e. Reject H_0.

11.27 Assuming that the error terms are independent and normally distributed with 0 mean and constant variance σ^2:

a. We know that $Z = \dfrac{\hat{\beta}_i - \beta_{i,0}}{\sigma\sqrt{c_{ii}}}$ has a standard normal distribution under H0.

Furthermore, $V = (n-2)S^2/\sigma^2$ has a chi–square distribution with n – 2 degrees of freedom. Therefore, by Definition 7.2,

$$\frac{Z}{\sqrt{V/(n-2)}} = \frac{\hat{\beta}_i - \beta_{i,0}}{S\sqrt{c_{ii}}}$$

has a t–distribution with n – 2 degrees of freedom under H_0 for i = 1, 2.

b. Using the pivotal quantity expressed above, the result follows from the material in Section 8.8.

11.28 Restricting to Ω_0, the likelihood function is

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

It is not difficult to verify that the MLEs for β_0 and σ^2 under the restricted space are $\overline{Y}$ and $\frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$ (respectively). The MLEs have already been found for the unrestricted space so that the LRT simplifies to

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \left(\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}\right)^{n/2} = \left(\frac{\text{SSE}}{S_{yy}}\right)^{n/2}.$$

So, we reject if $\lambda \leq k$, or equivalently if

$$\frac{S_{yy}}{\text{SSE}} \geq k^{-2/n} = k'.$$

Using the result from 11.15,

$$\frac{\text{SSE} + \hat{\beta}_1 S_{xy}}{\text{SSE}} = 1 + \frac{\hat{\beta}_1 S_{xy}}{\text{SSE}} = 1 + \frac{\hat{\beta}_1^2 S_{xx}}{(n-2)S} = 1 + \frac{T^2}{(n-2)}.$$

So, we see that λ is small whenever $T = \dfrac{\hat{\beta}_1}{S\sqrt{c_{11}}}$ is <u>large in magnitude</u>, where $c_{11} = \dfrac{1}{S_{xx}}$.

This is the usual t–test statistic, so the result has been proven.

11.29 Let $\hat{\beta}_1$ and $\hat{\gamma}_1$ be the least–squares estimators for the linear models $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ and $W_i = \gamma_0 + \gamma_1 c_i + \varepsilon_i$ as defined in the problem Then, we have that:

- $E(\hat{\beta}_1 - \hat{\gamma}_1) = \beta_1 - \gamma_1$
- $V(\hat{\beta}_1 - \hat{\gamma}_1) = \sigma^2\left(\frac{1}{S_{xx}} + \frac{1}{S_{cc}}\right)$, where $S_{cc} = \sum_{i=1}^{m}(c_i - \bar{c})^2$
- $\hat{\beta}_1 - \hat{\gamma}_1$ follows a normal distribution, so that under H_0, $\beta_1 - \gamma_1 = 0$ so that
$$Z = \frac{\hat{\beta}_1 - \hat{\gamma}_1}{\sigma\sqrt{\left(\frac{1}{S_{xx}} + \frac{1}{S_{cc}}\right)}} \text{ is standard normal}$$
- Let $V = \text{SSE}_Y + \text{SSE}_W = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 + \sum_{i=1}^{m}(W_i - \hat{W}_i)^2$. Then, V/σ^2 has a chi–square distribution with $n + m - 4$ degrees of freedom
- By Definition 7.2 we can build a random variable with a t–distribution (under H_0):
$$T = \frac{Z}{\sqrt{V/(n+m-4)}} = \frac{\hat{\beta}_1 - \hat{\gamma}_1}{S\sqrt{\left(\frac{1}{S_{xx}} + \frac{1}{S_{cc}}\right)}}, \text{ where } S = (\text{SSE}_Y + \text{SSE}_W)/(n + m - 4).$$

H_0 is rejected in favor of H_a for large values of $|T|$.

11.30 **a.** For the first experiment, the computed test statistic for H_0: $\beta_1 = 0$ vs. H_a: $\beta_1 \neq 0$ is $t_1 = (.155)/(.0202) = 7.67$ with 29 degrees of freedom. For the second experiment, the computed test statistic is $t_2 = (.190)/(.0193) = 9.84$ with 9 degrees of freedom. Both of these values reject the null hypothesis at $\alpha = .05$, so we can conclude that the slopes are significantly different from 0.

b. Using the result from Ex. 11.29, $S = (2.04 + 1.86)/(31 + 11 - 4) = .1026$. We can extract the values of S_{xx} and S_{cc} from the given values of $V(\hat{\beta}_1)$:
$$S_{xx} = \frac{SSE_Y/(n-2)}{V(\hat{\beta}_1)} = \frac{2.04/29}{(.0202)^2} = 172.397,$$
so similarly $S_{cc} = 554.825$. So, to test equality for the slope parameters, the computed test statistic is
$$|t| = \frac{|.155 - .190|}{\sqrt{.1024\left(\frac{1}{172.397} + \frac{1}{554.825}\right)}} = 1.25$$
with 38 degrees of freedom. Since $t_{.025} \approx z_{.025} = 1.96$, we fail to reject H_0: we cannot conclude that the slopes are different.

11.31 Here, R is used to fit the regression model:

```
> x <- c(19.1, 38.2, 57.3, 76.2, 95, 114, 131, 150, 170)
> y <- c(.095, .174, .256, .348, .429, .500, .580, .651, .722)
> summary(lm(y~x))

Call:
lm(formula = y ~ x)
```

```
Residuals:
       Min         1Q     Median         3Q        Max
-1.333e-02 -4.278e-03 -2.314e-05  8.056e-03  9.811e-03

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.875e-02  6.129e-03   3.059   0.0183 *
x           4.215e-03  5.771e-05  73.040 2.37e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.008376 on 7 degrees of freedom
Multiple R-Squared: 0.9987,     Adjusted R-squared: 0.9985
F-statistic:  5335 on 1 and 7 DF,  p-value: 2.372e-11
```

From the output, the fitted model is $\hat{y}$ = .01875 + .004215x. To test H_0: $\beta_1 = 0$ against H_a: $\beta_1 \neq 0$, note that the p–value is quite small indicating a very significant test statistic. Thus, H_0 is rejected and we can conclude that peak current increases as nickel concentrations increase (note that this is a one–sided alternative, so the p–value is actually `2.37e-11` divided by 2).

11.32 **a.** From Ex. 11.5, $\hat{\beta}_1$ = 4.8417 and S_{xx} = 42. From Ex. 11.15, s^2 = 3.0476 so to test H_0: $\beta_1 = 0$ vs. H_a: $\beta_1 > 0$, the required test statistic is t = 17.97 with 6 degrees of freedom. Since $t_{.01}$ = 3.143, H_0 is rejected: there is evidence of an increase.

b. The 99% CI for β_1 is 4.84 ± 1.00 or (3.84, 5.84).

11.33 Using the coded x's from 11.18, $\hat{\beta}_1^*$ = .655 and s^2 = 10.97. Since $S_{xx} = \sum_{i=1}^{10}\left(x_i^*\right)^2$ = 2360.2388, the computed test statistic is $|t| = \frac{|.655|}{\sqrt{\frac{10.97}{2360.2388}}}$ = 9.62 with 8 degrees of freedom. Since $t_{.025}$ = 2.306, we can conclude that there is evidence of a linear relationship.

11.34 **a.** Since $t_{.005}$ = 3.355, we have that p–value < 2(.005) = .01.
b. Using the Applet, $2P(T < 9.61)$ = 2(.00001) = .00002.

11.35 With $a_0 = 1$ and $a_1 = x^*$, the result follows since

$$V(\hat{\beta}_0 + \hat{\beta}_1 x^*) = \frac{1\cdot\frac{1}{n}\sum_{i=1}^n x_i^2 + (x^*)^2 - 2x^*\bar{x}}{S_{xx}}\sigma^2 = \frac{\frac{1}{n}\left(\sum_{i=1}^n x_i^2 - n\bar{x}^2\right) + (x^*)^2 - 2x^*\bar{x} + \bar{x}^2}{S_{xx}}\sigma^2$$

$$= \frac{\frac{1}{n}S_{xx} + (x^* - \bar{x})^2}{S_{xx}}\sigma^2 = \left[\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}\right]\sigma^2 .$$

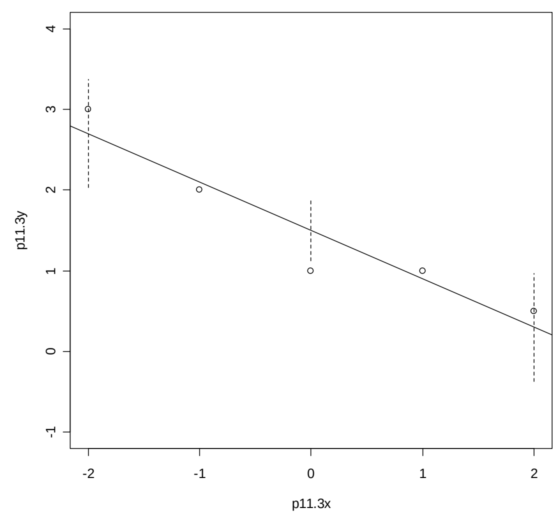
This is minimized when $(x^* - \bar{x})^2 = 0$, so $x^* = \bar{x}$.

11.36 From Ex. 11.13 and 11.24, when $x^* = 5$, $\hat{y} = 452.119 - 29.402(5) = 305.11$ so that $V(\hat{Y})$ is estimated to be 402.98. Thus, a 90% CI for $E(Y)$ is $305.11 \pm 1.782\sqrt{402.98}$ = 305.11 ± 35.773.

11.37 From Ex. 11.8 and 11.18, when $x^* = 12$, $\hat{y} = 7.15$ so that $V(\hat{Y})$ is estimated to be $10.97\left[.1 + \frac{(12-15.504)^2}{2359.929}\right] = 1.154$ Thus, a 95% CI for $E(Y)$ is $7.15 \pm 2.306\sqrt{1.154}$ = 7.15 ± 2.477 or (4.67, 9.63).

11.38 Refer to Ex. 11.3 and 11.23, where $s^2 = .1333$, $\hat{y} = 1.5 - .6x$, $S_{xx} = 10$ and $\bar{x} = 0$.

- When $x^* = 0$, the 90% CI for $E(Y)$ is $1.5 \pm 2.353\sqrt{.1333(\frac{1}{5})}$ or (1.12, 1.88).
- When $x^* = -2$, the 90% CI for $E(Y)$ is $2.7 \pm 2.353\sqrt{.1333(\frac{1}{5} + \frac{4}{10})}$ or (2.03, 3.37).
- When $x^* = 2$, the 90% CI for $E(Y)$ is $.3 \pm 2.353\sqrt{.1333(\frac{1}{5} + \frac{4}{10})}$ or (–.37, .97).



On the graph, note the interval lengths.

11.39 Refer to Ex. 11.16. When $x^* = 65$, $\hat{y} = 25.395$ and a 95% CI for $E(Y)$ is

$$25.395 \pm 2.228\sqrt{19.033\left[\frac{1}{12} + \frac{(65-60)^2}{6000}\right]}$$ or 25.395 ± 2.875.

11.40 Refer to Ex. 11.14. When $x^* = .3$, $\hat{y} = .7878$ and with SSE = .0155, $S_{xx} = .20625$, and $\bar{x} = .325$, the 90% CI for $E(Y)$ is $.7878 \pm 1.86\sqrt{\frac{.0155}{8}\left[\frac{1}{10} + \frac{(.3-.325)^2}{.20625}\right]}$ or (.76, .81).

11.41 **a.** Using $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{x}$ and $\hat{\beta}_1$ as estimators, we have $\hat{\mu}_y = \overline{Y} - \hat{\beta}_1\overline{x} + \hat{\beta}_1\mu_x$ so that

$$\hat{\mu}_y = \overline{Y} - \hat{\beta}_1(\overline{x} - \mu_x).$$

b. Calculate $V(\hat{\mu}_y) = V(\overline{Y}) + (\overline{x} - \mu_x)^2 V(\hat{\beta}_1) = \frac{\sigma^2}{n} + (\overline{x} - \mu_x)^2 \frac{\sigma^2}{S_{xx}} = \sigma^2\left(\frac{1}{n} + \frac{(\overline{x}-\mu_x)^2}{S_{xx}}\right)$.
From Ex. 11.4, $s^2 = 7.1057$ and $S_{xx} = 54{,}714$ so that $\hat{\mu}_y = 72.1 + .99(74 - 72) = 74.08$ and the variance of this estimate is calculated to be $7.1057\left[\frac{1}{10} + \frac{(74-72)^2}{54{,}714}\right] = .711$. The two–standard deviation error bound is $2\sqrt{.711} = 1.69$.

11.42 Similar to Ex. 11.35, the variance is minimized when $x^* = \overline{x}$.

11.43 Refer to Ex. 11.5 and 11.17. When $x = 9$ (year 1980), $\hat{y} = 65.15$ and the 95% PI is

$$65.15 \pm 2.447\sqrt{3.05\left(1 + \tfrac{1}{8} + \tfrac{(9-4.5)^2}{42}\right)} = 65.15 \pm 5.42 \text{ or } (59.73, 70.57).$$

11.44 For the year 1981, $x = 10$. So, $\hat{y} = 69.99$ and the 95% PI is

$$69.99 \pm 2.447\sqrt{3.05\left(1 + \tfrac{1}{8} + \tfrac{(10-4.5)^2}{42}\right)} = 69.99 \pm 5.80.$$

For the year 1982, $x = 11$. So, $\hat{y} = 74.83$ and the 95% PI is

$$74.83 \pm 2.447\sqrt{3.05\left(1 + \tfrac{1}{8} + \tfrac{(11-4.5)^2}{42}\right)} = 74.83 \pm 6.24.$$

Notice how the intervals get wider the further the prediction is from the mean. For the year 1988, this is far beyond the limits of experimentation. So, the linear relationship may not hold (note that the intervals for 1980, 1981 and 1982 are also outside of the limits, so caveat emptor).

11.45 From Ex. 11.8 and 11.18 (also see 11.37), when $x^* = 12$, $\hat{y} = 7.15$ so that the 95% PI is

$$7.15 \pm 2.306\sqrt{10.97\left[1 + \frac{1}{10} + \frac{(12 - 15.504)^2}{2359.929}\right]} = 7.15 \pm 8.03 \text{ or } (-.86, 15.18).$$

11.46 From 11.16 and 11.39, when $x^* = 65$, $\hat{y} = 25.395$ so that the 95% PI is given by

$$25.395 \pm 2.228\sqrt{19.033\left[1 + \frac{1}{12} + \frac{(65 - 60)^2}{6000}\right]} = 25.395 \pm 10.136.$$

11.47 From Ex. 11.14, when $x^* = .6$, $\hat{y} = .3933$ so that the 95% PI is given by

$$.3933 \pm 2.306\sqrt{.00194\left[1 + \frac{1}{10} + \frac{(.6 - .325)^2}{.20625}\right]} = .3933 \pm .12 \text{ or } (.27, .51).$$

11.48 The summary statistics are $S_{xx} = 380.5$, $S_{xy} = 2556.0$, and $S_{yy} = 19{,}263.6$. Thus, $r = .944$. To test H_0: $\rho = 0$ vs. H_a: $\rho > 0$, $t = 8.0923$ with 8 degrees of freedom. From Table 7, we find that p–value < .005.

11.49 **a.** r^2 behaves inversely to SSE, since $r^2 = 1 - \text{SSE}/S_{yy}$.
b. The best model has $r^2 = .817$, so $r = .90388$ (since the slope is positive, r is as well).

11.50 **a.** r^2 increases as the fit improves.
b. For the best model, $r^2 = .982$ and so $r = .99096$.
c. The scatterplot in this example exhibits a smaller error variance about the line.

11.51 The summary statistics are $S_{xx} = 2359.929$, $S_{xy} = 1546.459$, and $S_{yy} = 1101.1686$. Thus, $r = .9593$. To test H_0: $\rho = 0$ vs. H_a: $\rho \neq 0$, $|t| = 9.608$ with 8 degrees of freedom. From Table 7, we see that p–value < 2(.005) = .01 so we can reject the null hypothesis that the correlation is 0.

11.52 **a.** Since the slope of the line is negative, $r = -\sqrt{r^2} = -\sqrt{.61} = -.781$.
b. This is given by r^2, so 61%.
c. To test H_0: $\rho = 0$ vs. H_a: $\rho < 0$, $t = \frac{-.781\sqrt{12}}{\sqrt{1-(-.781)^2}} = -4.33$ with 12 degrees of freedom.
Since $-t_{.05} = -1.782$, we can reject H_0 and conclude that plant density decreases with increasing altitude.

11.53 **a.** This is given by $r^2 = (.8261)^2 = .68244$, or 68.244%.
b. Same answer as part **a**.
c. To test H_0: $\rho = 0$ vs. H_a: $\rho > 0$, $t = \frac{.8261\sqrt{8}}{\sqrt{1-(.8261)^2}} = 4.146$ with 8 degrees of freedom. Since $t_{.01} = 2.896$, we can reject H_0 and conclude that heights and weights are positively correlated for the football players.

d. p–value = $P(T > 4.146) = .00161$.

11.54 **a.** The MOM estimators for σ_X^2 and σ_Y^2 were given in Ex. 9.72.
b. By substituting the MOM estimators, the MOM estimator for ρ is identical to r, the MLE.

11.55 Since $\hat{\beta}_1 = S_{xy}/S_{xx}$ and $r = \hat{\beta}_1\sqrt{S_{xx}/S_{yy}}$, we have that the usual t–test statistic is:

$$T = \frac{\hat{\beta}_1}{\sqrt{S/S_{xx}}} = \frac{\sqrt{S_{xx}}\hat{\beta}_1\sqrt{n-2}}{\sqrt{S_{yy} - \hat{\beta}_1 S_{xy}}} = \frac{\sqrt{S_{xx}/S_{yy}}\hat{\beta}_1\sqrt{n-2}}{\sqrt{1-\hat{\beta}_1 S_{xy}/S_{yy}}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}.$$

11.56 Here, $r = .8$.
a. For $n = 5$, $t = 2.309$ with 3 degrees of freedom. Since $t_{.05} = 2.353$, fail to reject H_0.
b. For $n = 12$, $t = 4.2164$ with 10 degrees of freedom. Here, $t_{.05} = 1.812$, reject H_0.
c. For part a, p–value = $P(T > 2.309) = .05209$. For part (b), p–value = .00089.
d. Different conclusions: note the $\sqrt{n-2}$ term in the numerator of the test statistic.
e. The larger sample size in part b caused the computed test statistic to be more extreme. Also, the degrees of freedom were larger.

11.57 **a.** The sample correlation r determines the sign.
b. Both r and n determine the magnitude of $|t|$.

11.58 For the test H_0: $\rho = 0$ vs. H_a: $\rho > 0$, we reject if $t = \frac{r\sqrt{2}}{\sqrt{1-r^2}} \geq t_{.05} = 2.92$. The smallest value of r that would lead to a rejection of H_0 is the solution to the equation

$$r = \tfrac{2.92}{\sqrt{2}}\sqrt{1-r^2}\,.$$

Numerically, this is found to be $r = .9000$.

11.59 For the test H_0: $\rho = 0$ vs. H_a: $\rho < 0$, we reject if $t = \frac{r\sqrt{18}}{\sqrt{1-r^2}} \leq -t_{.05} = -1.734$. The largest value of r that would lead to a rejection of H_0 is the solution to the equation

$$r = \tfrac{-1.734}{\sqrt{18}}\sqrt{1-r^2}\,.$$

Numerically, this is found to be $r = -.3783$.

11.60 Recall the approximate normal distribution of $\frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)$ given on page 606. Therefore, for sample correlations r_1 and r_2, each being calculated from independent samples of size n_1 and n_2 (respectively) and drawn from bivariate normal populations with correlations coefficients ρ_1 and ρ_2 (respectively), we have that

$$Z = \frac{\frac{1}{2}\ln\left(\frac{1+r_1}{1-r_1}\right) - \frac{1}{2}\ln\left(\frac{1+r_2}{1-r_2}\right) - \left[\frac{1}{2}\ln\left(\frac{1+\rho_1}{1-\rho_1}\right) - \frac{1}{2}\ln\left(\frac{1+\rho_2}{1-\rho_2}\right)\right]}{\frac{1}{\sqrt{n_1-3}} + \frac{1}{\sqrt{n_2-3}}}$$

is approximately standard normal for large n_1 and n_2.
Thus, to test H_0: $\rho_1 = \rho_2$ vs. H_a: $\rho_1 \neq \rho_2$ with $r_1 = .9593$, $n_1 = 10$, $r_2 = .85$, $n_2 = 20$, the computed test statistic is

$$z = \frac{\frac{1}{2}\ln\left(\frac{1.9593}{.0407}\right) - \frac{1}{2}\ln\left(\frac{1.85}{.15}\right)}{\frac{1}{\sqrt{7}} + \frac{1}{\sqrt{17}}} = 1.52\,.$$

Since the rejection region is all values $|z| > 1.96$ for $\alpha = .05$, we fail to reject H_0.

11.61 Refer to Example 11.10 and the results given there. The 90% PI is

$$.979 \pm 2.132(.045)\sqrt{1 + \tfrac{1}{6} + \tfrac{(1.5-1.457)^2}{.234}} = .979 \pm .104 \text{ or } (.875, 1.083).$$

11.62 Using the calculations from Example 11.11, we have $r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = .9904$. The proportion of variation described is $r^2 = (.9904)^2 = .9809$.

11.63 **a.** Observe that $\ln E(Y) = \ln\alpha_0 - \alpha_1 x$. Thus, the logarithm of the expected value of Y is linearly related to x. So, we can use the linear model

$$w_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$$

where $w_i = \ln y_i$, $\beta_0 = \ln\alpha_0$ and $\beta_1 = -\alpha_1$. In the above, note that we are assuming an additive error term that is in effect after the transformation. Using the method of least squares, the summary statistics are:

$\bar{x}$ = 5.5, Σx^2 = 385, $\bar{w}$ = 3.5505, S_{xw} = –.7825, S_{xx} = 82.5, and S_{ww} = .008448.
Thus, $\hat{\beta}_1$ = –.0095, $\hat{\beta}_0$ = 3.603 and $\hat{\alpha}_1$ = –(–.0095) = .0095, $\hat{\alpha}_0$ = exp(3.603) = 36.70.
Therefore, the prediction equation is $\hat{y} = 36.70e^{-.0095x}$.

b. To find a CI for α_0, we first must find a CI for β_0 and then merely transform the endpoints of the interval. First, we calculate the SSE using SSE = $S_{ww} - \hat{\beta}_1 S_{xw}$ = .008448 – (–.0095)(–.782481) = .0010265 and so s^2 = (.0010265)/8 = .0001283 Using the methods given in Section 11.5, the 90% CI for β_0 is
$3.6027 \pm 1.86\sqrt{.0001283\left(\frac{385}{10(82.5)}\right)}$ or (3.5883, 3.6171). So the 90% CI for α_0 is given by

$$\left(e^{3.5883},\ e^{3.6171}\right) = (36.17, 37.23).$$

11.64 This is similar to Ex. 11.63. Note that $\ln E(Y) = -\alpha_0 x^{\alpha_1}$ and $\ln[-\ln E(Y)] = \ln\alpha_0 + \alpha_1 \ln x$. So, we would expect that ln(–ln$y$) to be linear in ln$x$. Define $w_i = \ln(-\ln y_i)$, $t_i = \ln x_i$, $\beta_0 = \ln\alpha_0$, $\beta_1 = \alpha_1$. So, we now have the familiar linear model

$$w_i = \beta_0 + \beta_1 t_i + \varepsilon_i$$

(again, we are assuming an additive error term that is in effect after the transformation). The methods of least squares can be used to estimate the parameters. The summary statistics are

$$\bar{t} = -1.12805,\ \bar{w} = -1.4616,\ S_{tw} = 3.6828, \text{ and } S_{tt} = 1.51548$$

So, $\hat{\beta}_1$ = 2.4142, $\hat{\beta}_0$ = 1.2617 and thus $\hat{\alpha}_1$ = 2.4142 and $\hat{\alpha}_0$ = exp(1.2617) = 3.5315.
This fitted model is $\hat{y} = \exp\left(-3.5315x^{2.4142}\right)$.

11.65 If y is related to t according to $y = 1 - e^{-\beta t}$, then $-\ln(1 - y) = \beta t$. Thus, let $w_i = -\ln(1 - y_i)$ and we have the linear model

$$w_i = \beta t_i + \varepsilon_i$$

(again assuming an additive error term). This is the "no–intercept" model described in Ex. 11.10 and the least squares estimator for β is given to be $\hat{\beta} = \frac{\sum_{i=1}^n t_i w_i}{\sum_{i=1}^n t_i^2}$. Now, using similar methods from Section 11.4, note that $V(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^n t_i^2}$ and $\frac{\text{SSE}}{\sigma^2} = \frac{\sum_{i=1}^n (w_i - \hat{w})^2}{\sigma^2}$ is chi–square with $n - 1$ degrees of freedom. So, by Definition 7.2, the quantity

$$T = \frac{\hat{\beta} - \beta}{S / \sum_{i=1}^n t_i^2},$$

where S = SSE/(n – 1) , has a t–distribution with n – 1 degrees of freedom.

A 100(1 – α)% CI for β is

$$\hat{\beta} \pm t_{\alpha/2} S \sqrt{\frac{1}{\sum_{i=1}^{n} t_i^2}},$$

and $t_{\alpha/2}$ is the upper– α/2 critical value from the t–distribution with $n - 1$ degrees of freedom.

11.66 Using the matrix notation from this section,

$$\boldsymbol{X} = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \quad \boldsymbol{Y} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \\ .5 \end{bmatrix} \quad \boldsymbol{X'Y} = \begin{bmatrix} 7.5 \\ -6 \end{bmatrix} \quad \boldsymbol{X'X} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}.$$

Thus, $\hat{\boldsymbol{\beta}} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}^{-1} \begin{bmatrix} 7.5 \\ -6 \end{bmatrix} = \begin{bmatrix} .2 & 0 \\ 0 & .1 \end{bmatrix} \begin{bmatrix} 7.5 \\ -6 \end{bmatrix} = \begin{bmatrix} 1.5 \\ -.6 \end{bmatrix}$ so that $\hat{y} = 1.5 - .6x$.

11.67 $$\boldsymbol{X} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \quad \boldsymbol{Y} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \\ .5 \end{bmatrix} \quad \boldsymbol{X'Y} = \begin{bmatrix} 7.5 \\ 1.5 \end{bmatrix} \quad \boldsymbol{X'X} = \begin{bmatrix} 5 & 5 \\ 5 & 15 \end{bmatrix}$$

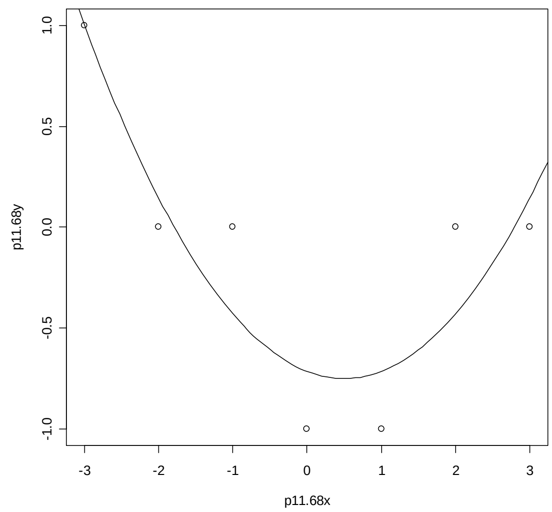
The student should verify that $(\boldsymbol{X'X})^{-1} = \begin{bmatrix} .3 & -.1 \\ -.1 & .1 \end{bmatrix}$ so that $\hat{\boldsymbol{\beta}} = \begin{bmatrix} 2.1 \\ -.6 \end{bmatrix}$. Not that the slope is the same as in Ex. 11.66, but the y–intercept is different. Since $\boldsymbol{X'X}$ is not a diagonal matrix (as in Ex. 11.66), computing the inverse is a bit more tedious.

11.68 $$\boldsymbol{X} = \begin{bmatrix} 1 & -3 & 9 \\ 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix} \quad \boldsymbol{Y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \quad \boldsymbol{X'Y} = \begin{bmatrix} -1 \\ 4 \\ 8 \end{bmatrix} \quad \boldsymbol{X'X} = \begin{bmatrix} 7 & 0 & 28 \\ 0 & 28 & 0 \\ 28 & 0 & 196 \end{bmatrix}.$$

The student should verify (either using Appendix I or a computer),

$$(\boldsymbol{X'X})^{-1} = \begin{bmatrix} .3333 & 0 & -.04762 \\ 0 & .035714 & 0 \\ -.04762 & 0 & .011905 \end{bmatrix} \text{ so that } \hat{\boldsymbol{\beta}} = \begin{bmatrix} -.714285 \\ -.142857 \\ .142859 \end{bmatrix} \text{ and the fitted}$$

model is $\hat{y} = -.714285 - .142857x + .142859x^2$.



The graphed curve is above.

11.69 For this problem, R will be used.

```
> x <- c(-7, -5, -3, -1, 1, 3, 5, 7)
> y <- c(18.5,22.6,27.2,31.2,33.0,44.9,49.4,35.0)
```

a. <u>Linear model</u>:

```
> lm(y~x)

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)            x
     32.725        1.812
```

$\leftarrow \hat{y} = 32.725 + 1.812x$

b. <u>Quadratic model</u>

```
> lm(y~x+I(x^2))

Call:
lm(formula = y ~ x + I(x^2))

Coefficients:
(Intercept)            x       I(x^2)
    35.5625       1.8119      -0.1351
```

$\leftarrow \hat{y} = 35.5625 + 1.8119x - .1351x^2$

11.70 **a.** The student should verify that $\boldsymbol{Y'Y} = 105{,}817$, $\boldsymbol{X'Y} = \begin{bmatrix} 721 \\ 106155 \end{bmatrix}$, and $\hat{\boldsymbol{\beta}} = \begin{bmatrix} .719805 \\ .991392 \end{bmatrix}$.

So, SSE = 105,817 – 105,760.155 = 56.845 and s^2 = 56.845/8 = 7.105625.

b. Using the coding as specified, the data are:

x_i^*	–62	–60	–63	–45	–25	40	–36	169	–13	95
y_i	9	14	7	29	45	109	40	238	60	70

The student should verify that $\boldsymbol{X}^{*'}\boldsymbol{Y} = \begin{bmatrix} 721 \\ 54243 \end{bmatrix}$, $\boldsymbol{X}^{*'}\boldsymbol{X}^{*} = \begin{bmatrix} 10 & 0 \\ 0 & 54{,}714 \end{bmatrix}$ and

$\hat{\boldsymbol{\beta}} = \begin{bmatrix} .72.1 \\ .991392 \end{bmatrix}$. So, SSE = 105,817 – 105,760.155 = 56.845 (same answer as part a).

11.71 Note that the vector **a** is composed of k 0's and one 1. Thus,

$$E(\hat{\beta}_i) = E(\mathbf{a}'\hat{\boldsymbol{\beta}}) = \mathbf{a}'E(\hat{\boldsymbol{\beta}}) = \mathbf{a}'\boldsymbol{\beta} = \beta_i$$

$$V(\hat{\beta}_i) = V(\mathbf{a}'\hat{\boldsymbol{\beta}}) = \mathbf{a}'E(\hat{\boldsymbol{\beta}})\mathbf{a} = \mathbf{a}'\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} = \sigma^2\mathbf{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} = c_{ii}\sigma^2$$

11.72 Following Ex. 11.69, more detail with the R output is given by:

```
> summary(lm(y~x+I(x^2)))

Call:
lm(formula = y ~ x + I(x^2))

Residuals:
     1      2      3      4      5      6      7      8
 2.242 -0.525 -1.711 -2.415 -4.239  5.118  8.156 -6.625

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  35.5625     3.1224  11.390 9.13e-05 ***
x             1.8119     0.4481   4.044  0.00988 **
I(x^2)       -0.1351     0.1120  -1.206  0.28167
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.808 on 5 degrees of freedom
Multiple R-Squared: 0.7808,     Adjusted R-squared: 0.6931
F-statistic: 8.904 on 2 and 5 DF,  p-value: 0.0225
```

a. To test H_0: $\beta_2 = 0$ vs. H_a: $\beta_2 \neq 0$, the computed test statistic is $t = –1.206$ and p–value = .28167. Thus, H_0 would not be rejected (no quadratic effect).

b. From the output, it is presented that $\sqrt{V(\hat{\beta}_2)}$ = .1120. So, with 5 degrees of freedom, $t_{.05} = 3.365$ so a 90% for β_2 is –.1351 ± (3.365)(.1120) = –.1351 ± .3769 or (–.512, .2418). Note that this interval contains 0, agreeing with part a.

11.73 If the minimum value is to occur at $x_0 = 1$, then this implies $\beta_1 + 2\beta_2 = 0$. To test this claim, let $\mathbf{a}' = [0\ \ 1\ \ 2]$ for the hypothesis H_0: $\beta_1 + 2\beta_2 = 0$ vs. H_a: $\beta_1 + 2\beta_2 \neq 0$. From

Ex. 11.68, we have that $\hat{\beta}_1 + 2\hat{\beta}_2 = .142861$, $s^2 = .14285$ and we calculate $\mathbf{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} =$ .083334. So, the computed value of the test statistic is $|t| = \frac{|.142861|}{\sqrt{.14285(.083334)}} = 1.31$ with 4 degrees of freedom. Since $t_{.025} = 2.776$, H_0 is not rejected.

11.74 **a.** Each transformation is defined, for each factor, by subtracting the midpoint (the mean) and dividing by one–half the range.

b. Using the matrix definitions of $\boldsymbol{X}$ and $\boldsymbol{Y}$, we have that

$$\boldsymbol{X}'\boldsymbol{Y} = \begin{bmatrix} 338 \\ -50.2 \\ -19.4 \\ -2.6 \\ -20.4 \end{bmatrix} \quad \boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} 16 & 0 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 & 0 \\ 0 & 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 & 16 \end{bmatrix} \text{ so that } \hat{\boldsymbol{\beta}} = \begin{bmatrix} 21.125 \\ -3.1375 \\ -1.2125 \\ -.1625 \\ -1.275 \end{bmatrix}.$$

The fitted model is $\hat{y} = 21.125 - 3.1375x_1 - 1.2125x_2 - .1625x_3 - 1.275x_4$.

c. First, note that SSE = $\boldsymbol{Y}'\boldsymbol{Y} - \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{Y}$ = 7446.52 – 7347.7075 = 98.8125 so that s^2 = 98.8125/(16 – 5) = 8.98. Further, tests of H_0: $\beta_i = 0$ vs. H_0: $\beta_i \neq 0$ for i = 1, 2, 3, 4, are based on the statistic $t_i = \dfrac{\hat{\beta}_i}{s\sqrt{c_{ii}}} = \dfrac{4\hat{\beta}_i}{\sqrt{8.98}}$ and H_0 is rejected if $|t_i| > t_{.005} = 3.106$. The four computed test statistics are $t_1 = -4.19$, $t_2 = -1.62$, $t_3 = -.22$ and $t_4 = -1.70$. Thus, only the first hypothesis involving the first temperature factor is significant.

11.75 With the four given factor levels, we have $\mathbf{a}' = [1\ \ -1\ \ 1\ \ -1\ \ 1]$ and so $\mathbf{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} =$ 5/16. The estimate of the mean of Y at this setting is

$$\hat{y} = 21.125 + 3.1375 - 1.2125 + .1625 - 1.275 = 21.9375$$

and the 90% confidence interval (based on 11 degrees of freedom) is

$$21.9375 \pm 1.796\sqrt{8.96}\sqrt{5/16} = 21.9375 \pm 3.01 \text{ or } (18.93,\ 24.95).$$

11.76 First, we calculate $s^2 = \text{SSE}/(n - k - 1) = 1107.01/11 = 100.637$.

a. To test H_0: $\beta_2 = 0$ vs. H_0: $\beta_2 < 0$, we use the t–test with $c_{22} = 8.1 \cdot 10^{-4}$:

$$t = \frac{-.92}{\sqrt{100.637(.00081)}} = -3.222.$$

With 11 degrees of freedom, $-t_{.05} = -1.796$ so we reject H_0: there is sufficient evidence that $\beta_2 < 0$.

b. (Similar to Ex. 11.75) With the three given levels, we have $\mathbf{a}' = [1\ \ 914\ \ 65\ \ 6]$ and so $\mathbf{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} = 92.76617$. The estimate of the mean of Y at this setting is

$$\hat{y} = 38.83 - .0092(914) - .92(65) + 11.56(6) = 39.9812$$

and the 95% CI based on 11 degrees of freedom is

$$39.9812 \pm 2.201\sqrt{100.637}\sqrt{92.76617} = 39.9812 \pm 212.664.$$

11.77 Following Ex. 11.76, the 95% PI is $39.9812 \pm 2.201\sqrt{100.637}\sqrt{93.76617} = 39.9812 \pm 213.807$.

11.78 From Ex. 11.69, the fitted model is $\hat{y} = 35.5625 + 1.8119x - .1351x^2$. For the year 2004, $x = 9$ and the predicted sales is $\hat{y} = 35.5625 + 1.8119(9) - .135(9^2) = 40.9346$. With we have $\mathbf{a}' = [1 \quad 9 \quad 81]$ and so $\mathbf{a}'(\boldsymbol{X'X})^{-1}\mathbf{a} = 1.94643$. The 98% PI for *Lexus* sales in 2004 is then

$$40.9346 \pm 3.365(5.808)\sqrt{1 + 1.94643} = 40.9346 \pm 33.5475.$$

11.79 For the given levels, $\hat{y} = 21.9375$, $\mathbf{a}'(\boldsymbol{X'X})^{-1}\mathbf{a} = .3135$, and $s^2 = 8.98$. The 90% PI based on 11 degrees of freedom is $21.9375 \pm 1.796\sqrt{8.98(1 + .3135)} = 21.9375 \pm 6.17$ or (15.77, 28.11).

11.80 Following Ex. 11.31, $S_{yy} = .3748$ and SSE = $S_{yy} - \hat{\beta}_1 S_{xy} = .3748 - (.004215)(88.8) = .000508$. Therefore, the F–test is given by $F = \frac{(.3748 - .000508)/1}{.000508/7} = 5157.57$ with 1 numerator and 7 denominator degrees of freedom. Clearly, p–value < .005 so reject H_0.

11.81 From Definition 7.2, let $Z \sim \text{Nor}(0, 1)$ and $W \sim \chi^2_\nu$, and let Z and W be independent. Then, $T = \dfrac{Z}{\sqrt{W/\nu}}$ has the t–distribution with ν degrees of freedom. But, since $Z^2 \sim \chi^2_1$, by Definition 7.3, $F = T^2$ has a F–distribution with 1 numerator and ν denominator degrees of freedom. Now, specific to this problem, note that if $k = 1$, $\text{SSE}_R = Syy$. So, the reduced model F–test simplifies to

$$F = \frac{S_{yy} - (S_{yy} - \hat{\beta}_1 S_{xy})}{SSE_C/(n-2)} = \frac{\hat{\beta}_1^2}{s^2/S_{xx}} = T^2.$$

11.82 **a.** To test H_0: $\beta_1 = \beta_2 = \beta_3 = 0$ vs. H_a: at least one $\beta_i \neq 0$, the F–statistic is

$$F = \frac{(10965.46 - 1107.01)/3}{1107.01/11} = 32.653,$$

with 3 numerator and 11 denominator degrees of freedom. From Table 7, we see that p–value < .005, so there is evidence that at least one predictor variable contributes.

b. The coefficient of determination is $R^2 = 1 - \dfrac{1107.01}{10965.46} = .899$, so 89.9% of the variation in percent yield (Y) is explained by the model.

11.83 **a.** To test H_0: $\beta_2 = \beta_3 = 0$ vs. H_a: at least one $\beta_i \neq 0$, the reduced model F–test is

$$F = \frac{(5470.07 - 1107.01)/2}{1107.01/11} = 21.677,$$

with 2 numerator and 11 denominator degrees of freedom. Since $F_{.05}$ = 3.98, we can reject H_0.

b. We must find the value of SSE_R such that $\dfrac{(\text{SSE}_R - 1107.01)/2}{1107.01/11}$ = 3.98. The solution is SSE_R = 1908.08

11.84 **a.** The result follows from

$$\frac{n-(k+1)}{k}\left(\frac{R^2}{1-R^2}\right) = \frac{n-(k+1)}{k}\left(\frac{1-\text{SSE}/S_{yy}}{\text{SSE}/S_{yy}}\right) = \frac{(S_{yy} - \text{SSE})/k}{\text{SSE}/[n-(k+1)]} = F.$$

b. The form is $F = T^2$.

11.85 Here, n = 15, k = 4.

a. Using the result from Ex. 11.84, $F = \dfrac{10}{4}\left(\dfrac{.942}{1-.942}\right)$ = 40.603 with 4 numerator and 10 denominator degrees of freedom. From Table 7, it is clear that p–value < .005, so we can safely conclude that at least one of the variables contributes to predicting the selling price.

b. Since $R^2 = 1 - \text{SSE}/S_{yy}$, SSE = 16382.2(1 – .942) = 950.1676.

11.86 To test H_0: $\beta_2 = \beta_3 = \beta_4 = 0$ vs. H_a: at least one $\beta_i \neq 0$, the reduced–model F–test is

$$F = \frac{(1553 - 950.16)/3}{950.1676/10} = 2.115,$$

with 3 numerator and 10 denominator degrees of freedom. Since $F_{.05}$ = 3.71, we fail to reject H_0 and conclude that these variables should be dropped from the model.

11.87 **a.** The F–statistic, using the result in Ex. 11.84, is $F = \dfrac{2}{4}\left(\dfrac{.9}{.1}\right)$ = 4.5 with 4 numerator and 2 denominator degrees of freedom. Since $F_{.1}$ = 9.24, we fail to reject H_0.

b. Since k is large with respect to n, this makes the computed F–statistic small.

c. The F–statistic, using the result in Ex. 11.84, is $F = \dfrac{40}{3}\left(\dfrac{.15}{.85}\right)$ = 2.353 with 3 numerator and 40 denominator degrees of freedom. Since $F_{.1}$ = 2.23, we can reject H_0.

d. Since k is small with respect to n, this makes the computed F–statistic large.

11.88 **a.** False; there are 15 degrees of freedom for SSE.
b. False; the fit (R^2) cannot improve when independent variables are removed.
c. True

d. False; not necessarily, since the degrees of freedom associated with each SSE is different.
e. True.
f. False; Model III is not a reduction of Model I (note the x_1x_2 term).

11.89 **a.** True.
b. False; not necessarily, since Model III is not a reduction of Model I (note the x_1x_2 term).
c. False; for the same reason in part (b).

11.90 Refer to Ex. 11.69 and 11.72.
a. We have that $\text{SSE}_R = 217.7112$ and $\text{SSE}_C = 168.636$. For H_0: $\beta_2 = 0$ vs. H_a: $\beta_2 \neq 0$, the reduced model F–test is $F = \dfrac{217.7112 - 168.636}{168.636/5} = 1.455$ with 1 numerator and 5 denominator degrees of freedom. With $F_{.05} = 6.61$, we fail to reject H_0.

b. Referring to the R output given in Ex. 11.72, the F–statistic is $F = 8.904$ and the p–value for the test is .0225. This leads to a rejection at the $\alpha = .05$ level.

11.91 The hypothesis of interest is H_0: $\beta_1 = \beta_4 = 0$ vs. H_a: at least one $\beta_i \neq 0$, $i = 1, 4$. From Ex. 11.74, we have $\text{SSE}_C = 98.8125$. To find SSE_R, we fit the linear regression model with just x_2 and x_3 so that

$$\boldsymbol{X'Y} = \begin{bmatrix} 338 \\ -19.4 \\ -2.6 \end{bmatrix} \qquad (\boldsymbol{X'X})^{-1} = \begin{bmatrix} 1/16 & 0 & 0 \\ 0 & 1/16 & 0 \\ 0 & 0 & 1/16 \end{bmatrix}$$

and so $\text{SSE}_R = 7446.52 - 7164.195 = 282.325$. The reduced–model F–test is

$$F = \frac{(282.325 - 98.8125)/2}{98.8125/11} = 10.21,$$

with 2 numerator and 11 denominator degrees of freedom. Thus, since $F_{.05} = 3.98$, we can reject H_0 can conclude that either T_1 or T_2 (or both) affect the yield.

11.92 To test H_0: $\beta_3 = \beta_4 = \beta_5 = 0$ vs. H_a: at least one $\beta_i \neq 0$, the reduced–model F–test is

$$F = \frac{(465.134 - 152.177)/3}{152.177/18} = 12.34,$$

with 3 numerator and 18 denominator degrees of freedom. Since $F_{.005} = 5.92$, we have that p–value < .005.

11.93 Refer to Example. 11.19. For the reduced model, $s^2 = 326.623/8 = 40.83$. Then,

$$(\boldsymbol{X'X})^{-1} = \begin{bmatrix} 1/11 & 0 & 0 \\ 0 & 2/17 & 0 \\ 0 & 0 & 2/17 \end{bmatrix}, \ \mathbf{a'} = [1 \quad 1 \quad -1].$$

So, $\hat{y} = \mathbf{a}'\hat{\boldsymbol{\beta}}$ = 93.73 + 4 – 7.35 = 90.38 and $\mathbf{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} = .3262$. The 95% CI for $E(Y)$ is $90.38 \pm 2.306\sqrt{40.83(.3262)}$ = 90.38 ± 8.42 or (81.96, 98.80).

11.94 From Example 11.19, tests of H_0: $\beta_i = 0$ vs. H_0: $\beta_i \neq 0$ for i = 3, 4, 5, are based on the statistic $t_i = \dfrac{\hat{\beta}_i}{s\sqrt{c_{ii}}}$ with 5 degrees of freedom and H_0 is rejected if $|t_i| > t_{.01} = 4.032$.
The three computed test statistics are $|t_3| = .58$, $|t_4| = 3.05$, $|t_5| = 2.53$. Therefore, none of the three parameters are significantly different from 0.

11.95 **a.** The summary statistics are: $\bar{x}$ = –268.28, $\bar{y}$ = .6826, S_{xy} = –15.728, S_{xx} = 297.716, and S_{yy} = .9732. Thus, $\hat{y}$ = –13.54 – 0.053x.
b. First, SSE = .9732 – (–.053)(–15.728) = .14225, so s^2 = .14225/8 = .01778. The test statistic is $t = \frac{-.053}{\sqrt{\frac{.01778}{297.716}}}$ = –6.86 and H_0 is rejected at the α = .01 level.
c. With x = –273, $\hat{y}$ = –13.54 – .053(–273) = .929. The 95% PI is
$.929 \pm 2.306\sqrt{.01778}\sqrt{1 + \frac{1}{10} + \frac{(273-268.28)^2}{297.716}}$ = .929 ± .33.

11.96 Here, R will be used to fit the model:

```
> x <- c(.499, .558, .604, .441, .550, .528, .418, .480, .406, .467)
> y <- c(11.14,12.74,13.13,11.51,12.38,12.60,11.13,11.70,11.02,11.41)
> summary(lm(y~x))

Call:
lm(formula = y ~ x)

Residuals:
     Min       1Q   Median       3Q      Max
-0.77823 -0.07102  0.08181  0.16435  0.36771

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   6.5143     0.8528   7.639 6.08e-05 ***
x            10.8294     1.7093   6.336 0.000224 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3321 on 8 degrees of freedom
Multiple R-Squared: 0.8338,     Adjusted R-squared: 0.813
F-statistic: 40.14 on 1 and 8 DF,  p-value: 0.0002241
```

a. The fitted model is $\hat{y}$ = 6.5143 + 10.8294x.
b. The test H_0: $\beta_1 = 0$ vs. H_a: $\beta_1 \neq 0$ has a p–value of .000224, so H_0 is rejected.
c. It is found that s = .3321 and S_{xx} = .0378. So, with x = .59, $\hat{y}$ = 6.5143 + 10.8294(.59) = 12.902. The 90% CI for $E(Y)$ is
$12.902 \pm 1.860(.3321)\sqrt{\frac{1}{10} + \frac{(.59-.4951)^2}{.0378}}$ = 12.902 ± .36.

11.97 **a.** Using the matrix notation,

$$\boldsymbol{X} = \begin{bmatrix} 1 & -3 & 5 & -1 \\ 1 & -2 & 0 & 1 \\ 1 & -1 & -3 & 1 \\ 1 & 0 & -4 & 0 \\ 1 & 1 & -3 & -1 \\ 1 & 2 & 0 & -1 \\ 1 & 3 & 5 & 1 \end{bmatrix}, \ \boldsymbol{Y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}, \ \boldsymbol{X'Y} = \begin{bmatrix} 10 \\ 14 \\ 10 \\ -3 \end{bmatrix}, \ (\boldsymbol{X'X})^{-1} = \begin{bmatrix} 1/7 & 0 & 0 & 0 \\ 0 & 1/28 & 0 & 0 \\ 0 & 0 & 1/84 & 0 \\ 0 & 0 & 0 & 1/6 \end{bmatrix}.$$

So, the fitted model is found to be $\hat{y}$ = 1.4825 + .5x_1 + .1190x_2 – .5x_3.

b. The predicted value is $\hat{y}$ = 1.4825 + .5 – .357 + .5 = 2.0715. The observed value at these levels was y = 2. The predicted value was based on a model fit (using all of the data) and the latter is an observed response.

c. First, note that SSE = 24 – 23.9757 = .0243 so s^2 = .0243/3 = .008. The test statistic is $t = \frac{\hat{\beta}_3}{s\sqrt{c_{ii}}} = \frac{-.5}{\sqrt{.008(1/6)}}$ = –13.7 which leads to a rejection of the null hypothesis.

d. Here, $\mathbf{a}'$ = [1 1 –3 –1] and so $\mathbf{a}'(\boldsymbol{X'X})^{-1}\mathbf{a}$ = .45238. So, the 95% CI for $E(Y)$ is $2.0715 \pm 3.182\sqrt{.008}\sqrt{.45238}$ = 2.0715 ± .19 or (1.88, 2.26).

e. The prediction interval is $2.0715 \pm 3.182\sqrt{.008}\sqrt{1+.45238}$ = 2.0715 ± .34 or (1.73, 2.41).

11.98 Symmetric spacing about the origin creates a diagonal **X′X** matrix which is very easy to invert.

11.99 Since $V(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$, this will be minimized when $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$ is as large as possible. This occurs when the x_i are as far away from $\bar{x}$ as possible. If $-9 \le x \le 9$, chose n/2 at x = –9 and n/2 at x = 9.

11.100 Based on the minimization strategy in Ex. 11.99, the values of x are: –9, –9, –9, –9, –9, 9, 9, 9, 9, 9. Thus $S_{xx} = \sum_{i=1}^{10}(x_i - \bar{x})^2 = \sum_{i=1}^{10} x_i^2$ = 810. If equal spacing is employed, the values of x are: –9, –7, –5, –3, –1, 1, 3, 5, 7, 9. Thus, $S_{xx} = \sum_{i=1}^{10}(x_i - \bar{x})^2 = \sum_{i=1}^{10} x_i^2$ = 330. The relative efficiency is the ratio of the variances, or 330/810 = 11/27.

11.101 Here, R will be used to fit the model:

```
> x1 <- c(0,0,0,0,0,1,1,1,1,1)
> x2 <- c(-2,-1,0,1,2,-2,-1,0,1,2)
> y <- c(8,9,9.1,10.2,10.4,10,10.3,12.2,12.6,13.9)
```

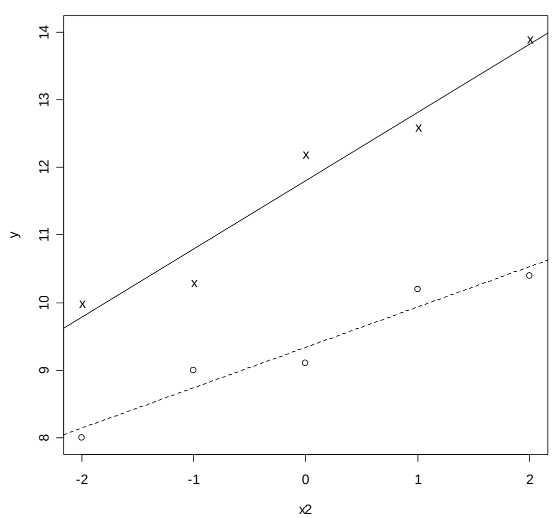
```
> summary(lm(y~x1+x2+I(x1*x2)))
Call:
lm(formula = y ~ x1 + x2 + I(x1 * x2))

Residuals:
    Min      1Q  Median      3Q     Max 
-0.4900 -0.1925 -0.0300  0.2500  0.4000 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   9.3400     0.1561  59.834 1.46e-09 ***
x1            2.4600     0.2208  11.144 3.11e-05 ***
x2            0.6000     0.1104   5.436  0.00161 ** 
I(x1 * x2)    0.4100     0.1561   2.627  0.03924 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
Residual standard error: 0.349 on 6 degrees of freedom
Multiple R-Squared: 0.9754,     Adjusted R-squared: 0.963 
F-statistic: 79.15 on 3 and 6 DF,  p-value: 3.244e-05
```

a. The fitted model is $\hat{y} = 9.34 + 2.46x_1 + .6x_2 + .41x_1x_2$.

b. For bacteria type A, $x_1 = 0$ so $\hat{y} = 9.34 + .6x_2$ (dotted line)
For bacteria type B, $x_1 = 1$ so $\hat{y} = 11.80 + 1.01\,x_2$ (solid line)



c. For bacteria A, $x_1 = 0$, $x_2 = 0$, so $\hat{y} = 9.34$. For bacteria B, $x_1 = 1$, $x_2 = 0$, so $\hat{y} = 11.80$. The observed growths were 9.1 and 12.2, respectively.

d. The rates are different if the parameter β_3 is nonzero. So, H_0: $\beta_3 = 0$ vs. H_a: $\beta_3 \neq 0$ has a p–value = .03924 (R output above) and H_0 is rejected.

e. With $x_1 = 1$, $x_2 = 1$, so $\hat{y} = 12.81$. With $s = .349$ and $\mathbf{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} = .3$, the 90% CI is $12.81 \pm .37$.

f. The 90% PI is $12.81 \pm .78$.

11.102 The reduced model F statistic is $F = \dfrac{(795.23 - 783.9)/2}{783.9/195} = 1.41$ with 2 numerator and 195 denominator degrees of freedom. Since $F_{.05} \approx 3.00$, we fail to reject H_0: salary is not dependent on gender.

11.103 Define $\boldsymbol{1}$ as a column vector of n 1's. Then $\bar{y} = \frac{1}{n}\boldsymbol{1}'\boldsymbol{Y}$. We must solve for the vector $\boldsymbol{x}$ such that $\bar{y} = \boldsymbol{x}'\hat{\boldsymbol{\beta}}$. Using the matrix definition of $\hat{\boldsymbol{\beta}}$, we have

$$\bar{y} = \boldsymbol{x}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} = \tfrac{1}{n}\boldsymbol{1}'\boldsymbol{Y}$$
$$\boldsymbol{x}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}\boldsymbol{Y}' = \tfrac{1}{n}\boldsymbol{1}'\boldsymbol{Y}\boldsymbol{Y}'$$

which implies

$$\boldsymbol{x}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' = \tfrac{1}{n}\boldsymbol{1}'$$
$$\boldsymbol{x}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X} = \tfrac{1}{n}\boldsymbol{1}'\boldsymbol{X}$$

so that

$$\boldsymbol{x}' = \tfrac{1}{n}\boldsymbol{1}'\boldsymbol{X}.$$

That is, $\boldsymbol{x}' = [1 \;\; \bar{x}_1 \;\; \bar{x}_2 \;\; \ldots \;\; \bar{x}_k]$.

11.104 Here, we will use the coding $x_1 = \frac{P-65}{15}$ and $x_2 = \frac{T-200}{100}$. Then, the levels are $x_1 = -1, 1$ and $x_2 = -1, 0, 1$.

a. $\boldsymbol{Y} = \begin{bmatrix} 21 \\ 23 \\ 26 \\ 22 \\ 23 \\ 28 \end{bmatrix}$ $\boldsymbol{X} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & -1 & 1 \end{bmatrix}$ $\boldsymbol{X}'\boldsymbol{Y} = \begin{bmatrix} 143 \\ 3 \\ 11 \\ 97 \end{bmatrix}$ $(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} .5 & 0 & 0 & -.5 \\ 0 & .1667 & 0 & 0 \\ 0 & 0 & .25 & 0 \\ -.5 & 0 & 0 & .75 \end{bmatrix}$

So, the fitted model is $\hat{y} = 23 + .5x_1 + 2.75x_2 + 1.25x_2^2$.

b. The hypothesis of interest is H_0: $\beta_3 = 0$ vs. H_a: $\beta_3 \neq 0$ and the test statistic is (verify that SSE = 1 so that $s^2 = .5$) $|t| = \frac{|1.25|}{\sqrt{.5(.75)}} = 2.040$ with 2 degrees of freedom. Since $t_{.025} = 4.303$, we fail to reject H_0.

c. To test H_0: $\beta_2 = \beta_3 = 0$ vs. H_a: at least one $\beta_i \neq 0$, $i = 2, 3$, the reduced model must be fitted. It can be verified that $SSE_R = 33.33$ so that the reduced model F–test is $F = 32.33$ with 2 numerator and 2 denominator degrees of freedom. It is easily seen that H_0 should be rejected; temperature does affect yield.

11.105 **a.** $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}\sqrt{\frac{S_{yy}}{S_{xx}}} = r\sqrt{\frac{S_{yy}}{S_{xx}}}$.

b. The conditional distribution of Y_i, given $X_i = x_i$, is (see Chapter 5) normal with mean $\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x_i - \mu_x)$ and variance $\sigma_y^2(1-\rho^2)$. Redefine $\beta_1 = \rho\frac{\sigma_y}{\sigma_x}$, $\beta_0 = \mu_y - \beta_1\mu_x$. So, if $\rho = 0$, $\beta_1 = 0$. So, using the usual t–statistic to test $\beta_1 = 0$, we have

$$T = \frac{\hat{\beta}_1}{S/\sqrt{S_{xx}}} = \frac{\hat{\beta}_1}{\sqrt{\frac{\text{SSE}}{n-2}}\sqrt{\frac{1}{S_{xx}}}} = \frac{\hat{\beta}_1\sqrt{(n-2)S_{xx}}}{\sqrt{(1-r^2)S_{yy}}}.$$

c. By part a, $\hat{\beta}_1 = r\sqrt{\frac{S_{yy}}{S_{xx}}}$ and the statistic has the form as shown. Note that the distribution only depends on $n-2$ and not the particular value x_i. So, the distribution is the same unconditionally.

11.106 The summary statistics are $S_{xx} = 66.54$, $S_{xy} = 71.12$, and $S_{yy} = 93.979$. Thus, $r = .8994$. To test H_0: $\rho = 0$ vs. H_a: $\rho \neq 0$, $|t| = 5.04$ with 6 degrees of freedom. From Table 7, we see that p–value < 2(.005) = .01 so we can reject the null hypothesis that the correlation is 0.

11.107 The summary statistics are $S_{xx} = 153.875$, $S_{xy} = 12.8$, and $S_{yy} = 1.34$.

a. Thus, $r = .89$.

b. To test H_0: $\rho = 0$ vs. H_a: $\rho \neq 0$, $|t| = 4.78$ with 6 degrees of freedom. From Table 7, we see that p–value < 2(.005) = .01 so we can reject the null hypothesis that the correlation is 0.

11.108 **a.-c.** Answers vary.