Solution for Section 7.1
Due May 03, 07

1. \( \bar{x} = \bar{y} = 4, \sum_{i=1}^{n} (x_i - \bar{x})^2 = 28, \sum_{i=1}^{n} (y_i - \bar{y})^2 = 28, \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = 23. \)

\[
\begin{align*}
r & = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \\
& = 0.8214.
\end{align*}
\]

2. (a) \( y \) has been replaced with \( y + 3, \) \( x \) is unchanged. This involves only adding a constant, which does not change the correlation coefficient.

(b) \( x \) has been replaced with \( 10x + 1, \) \( y \) has been replaced with \( y + 3. \) This involves only adding a constant and multiplying by a constant, neither of which changes the correlation coefficient.

(c) \( x \) has been replaced with \( 10y + 33, \) \( y \) has been replaced with \( 2x + 2. \) This involves only adding a constant, multiplying by a constant, and interchanging \( x \) and \( y; \) none of which changes the correlation coefficient.

3. (a) The correlation coefficient is appropriate. The points are approximately clustered around a line.

(b) The correlation coefficient is not appropriate. The relationship is curved, not linear.

(c) The correlation coefficient is not appropriate. The plot contains outliers.

9. (a) \( \bar{x} = 105.63, \bar{y} = 38.84, \sum_{i=1}^{n} (x_i - \bar{x})^2 = 2.821, \sum_{i=1}^{n} (y_i - \bar{y})^2 = 47.144, \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = -5.902, n = 10. \)

\[
\begin{align*}
r & = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \\
& = -0.51178.
\end{align*}
\]

\[
W = \frac{1}{2} \ln \frac{1+r}{1-r} = -0.56514, \sigma_W = \sqrt{1/(10-3)} = 0.377964.
\]

A 95% confidence interval for \( \mu_Y \) is \( W \pm 1.96\sigma_W \), or \((-1.30595, 0.17567). \)

A 95% confidence interval for \( r \) is \( \left( \frac{e^{2(1.30595)} - 1}{e^{2(1.30595)} + 1} \right), \) or \((-0.8632, 0.1739). \)
(b) The null and alternate hypotheses are $H_0: \rho \geq 0.3$ versus $H_1: \rho < 0.3$.

$$r = -0.51178, \quad W = \frac{1}{2} \ln \frac{1+r}{1-r} = -0.56514, \quad \sigma_W = \sqrt{1/(10-3)} = 0.377964.$$  

Under $H_0$, take $\rho = 0.3$, so $\sigma_W = \frac{1}{2} \ln \frac{1+0.3}{1-0.3} = 0.30952$.

The null distribution of $W$ is therefore normal with mean 0.30952 and standard deviation 0.377964.

The $z$-score of -0.56514 is $(-0.56514 - 0.30952)/0.377964 = -2.31$. Since the alternate hypothesis is of the form $\rho < \rho_0$, the $P$-value is the area to the left of $z = -2.31$.

Thus $P = 0.0104$.

We conclude that $\rho < 0.3$.

---

(c) $r = -0.51178, n = 10, U = r\sqrt{n-2}/\sqrt{1-r^2} = -1.6849$.

Under $H_0$, $U$ has a Student's $t$ distribution with $10 - 2 = 8$ degrees of freedom.

Since the alternate hypothesis is of the form $\rho \neq \rho_0$, the $P$-value is the sum of the areas to the right of $t = 1.6849$ and to the left of $t = -1.6849$.

From the $t$ table, 0.10 < $P$ < 0.20. A computer package gives $P = 0.130$.

It is plausible that $\rho = 0$.

---

10. (a) $r = 0.10, n = 400, U = r\sqrt{n-2}/\sqrt{1-r^2} = 2.00304$.

Under $H_0$, $U$ has a Student's $t$ distribution with $400 - 2 = 398$ degrees of freedom. Since the number of degrees of freedom is large, use the $z$ table.

Since the alternate hypothesis is $H_1: \rho > 0$, the $P$-value is the area to the right of $z = 2.01$.

Thus $P = 0.0222$.

We conclude that $\rho > 0$.

(b) No, we can only conclude that $\rho > 0$. We cannot conclude that $\rho$ is large.
Solution for Section 7.2

Due May 03, 07

1. (a) $245.82 + 1.13(65) = 319.27$ pounds

(b) The difference in $y$ predicted from a one-unit change in $x$ is the slope $\hat{\beta}_1 = 1.13$. Therefore the change in the number of lbs of steam predicted from a change of 5°C is $1.13(5) = 5.65$ pounds.

3. $r^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = 1 - \frac{1450}{9615} = 0.8492$.

5. (a) $-0.2967 + 0.2738(70) = 18.869$ in.

(b) Let $x$ be the required height. Then $19 = -0.2967 + 0.2738x$, so $x = 70.477$ in.

(c) No, some of the men whose points lie below the least-squares line will have shorter arms.

6. $n = 40$, $\sum_{i=1}^{n}(x_i - \bar{x})^2 = 98,775$, $\sum_{i=1}^{n}(y_i - \bar{y})^2 = 19.10$, $\bar{x} = 26.36$, $\bar{y} = 0.5188$.

$\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = 826.94$.

(a) $r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} = 0.602052$.

(b) The error sum of squares is $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = (1 - r^2) \sum_{i=1}^{n}(y_i - \bar{y})^2 = 12.177$.

The regression sum of squares is $\sum_{i=1}^{n}(y_i - \bar{y})^2 - \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = 19.10 - 12.177 = 6.923$.

The total sum of squares is $\sum_{i=1}^{n}(y_i - \bar{y})^2 = 19.10$.

(c) $\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = 0.008371956$ and $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 0.298115$.

The equation of the least-squares line is $y = 0.298115 + 0.008371956x$.

(d) $0.298115 + 0.008371956(40) = 0.633$ mm

(e) Let $x$ be the required temperature. Then $0.5 = 0.298115 + 0.008371956x$, so $x = 24.1^\circ$C.

(f) No. If the actual amount of warping happens to come out above the value predicted by the least-squares line, it could exceed 0.5 mm.
7. (a) The linear model is appropriate.

(b) \( \bar{x} = 19.5, \quad \bar{y} = 5.534286, \quad \sum_{i=1}^{n} (x_i - \bar{x})^2 = 368.125, \quad \sum_{i=1}^{n} (y_i - \bar{y})^2 = 9.928171, \quad \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = -57.1075. \)

\[ \hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} = -0.1551307 \text{ and } \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 8.55933. \]

The equation of the least-squares line is \( y = 8.55933 - 0.1551307x. \)

(c) By \( 0.1551307(5) = 0.776 \text{ miles per gallon}. \)

(d) \( 8.55933 - 0.1551307(15) = 6.23 \text{ miles per gallon}. \)

(e) miles per gallon per ton

(f) miles per gallon