# A

# Answers to Selected Problems and Exercises

**1.6.1** **a** $\hat{\mu}_Y = 6.989$. $6.140 \le \mu_Y \le 7.838$; the width is 1.698.
**b** $C[6.140 \le \mu_Y \le 7.838] = 0.80$.
**c** A 95% lower confidence bound for $\mu_Y$ is 5.889.
**d** $C[5.889 \le \mu_Y] = 0.95$.
**g** NH is not rejected at the $\alpha = 0.05$ level.
**h** $C[2.93 \le \sigma_Y \le 4.54] = 0.90$. Accept NH in (*g*).
**i** A 99% two-sided confidence interval for $\mu_Y$ is [5.205, 8.773]. The width is 3.568.

**1.7.1** **a** Not defined. **b** $f(16)/f(36) = 0.188386$. **c** Not defined.
**d** $f(34) + f(13) = 7563.3096$. **e** Not defined.

**1.7.2** **a** $\mu_Y(6, 1) = \beta_0 + 6\beta_1 + \beta_2$. **b** $\mu_Y(15, -4) = \beta_0 + 15\beta_1 - 64\beta_2$.
**c** Yes, no, no.

**1.8.1** **a** $A^T = \begin{bmatrix} 9 & 4 & 3 \\ 4 & 16 & 8 \\ 3 & 8 & 12 \end{bmatrix}$

**b** $C + B^T$ is defined.
$B + C$ is not defined.
$B + C^T$ is defined.
$AC$ is defined.
$CA$ is not defined.
$B - C^T$ is defined.

**c** $C + B^T = \begin{bmatrix} 24 & 27 & 22 & 28 \\ 38 & 59 & 32 & 22 \\ 34 & 35 & 62 & 58 \end{bmatrix}$

$B + C^T = \begin{bmatrix} 24 & 38 & 34 \\ 27 & 59 & 35 \\ 22 & 32 & 62 \\ 28 & 22 & 58 \end{bmatrix}$

$AC = \begin{bmatrix} 297 & 409 & 352 & 359 \\ 680 & 780 & 700 & 624 \\ 600 & 653 & 695 & 550 \end{bmatrix}$

$$B - C^T = \begin{bmatrix} 0 & -10 & -28 \\ -19 & 3 & -25 \\ -12 & -6 & -20 \\ -16 & -18 & 4 \end{bmatrix}$$

**1.8.2** $\hat{\beta} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

**1.8.3** **a** $X^TX = \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix}$

**b** $X^Ty = \begin{bmatrix} 12 \\ 31 \end{bmatrix}$

**c** $(X^TX)^{-1} = \begin{bmatrix} 1.5 & -0.5 \\ -0.5 & 0.2 \end{bmatrix}$

**d** $\hat{\beta} = \begin{bmatrix} 2.5 \\ 0.2 \end{bmatrix}$

**1.10.1** **Target population of items**: The collection of all state employees last year.
**Target population of numbers**: The collection of numbers of days of sick leave each state employee took last year.
**Study population of items and of numbers** can be defined to be the same as the target population of items and of numbers for this problem. The **parameter** to be studied is the average number of days of sick leave taken by state employees last year.

**1.10.3** **Target population of items**: The collection of all bottles of aspirin that will be manufactured next month by the company.
**Target population of numbers**: The response variable in this problem is an attribute variable, viz., damaged or undamaged. We assign a score of 1 for bottles that arrive damaged at the retail store and a score of 0 for undamaged bottles. In this way we can convert attribute data to numerical data. The target population of numbers is the set of scores (0 or 1) assigned to each bottle of aspirin to be manufactured next month.
**Study population of items**: Set of all aspirin bottles manufactured by the company last month.
**Study population of numbers**: Set of scores (0 or 1) of all bottles of aspirin manufactured by the company last month.
**Parameter of interest**: The proportion of bottles to be manufactured next month that will arrive damaged at the retail store. Note that this proportion is in fact the average of the 0, 1 scores assigned to the aspirin bottles!

**1.10.6** The proportion of the population that is less than 17.5 is 0.8554. The proportion of the population between 12 and 16 is equal to 0.4977. The proportion of the population that is greater than 15 is equal to 0.7611.

**1.10.7** **Target population of items**: All hospital patients who will stay at that hospital next year.
**Target population of numbers**: The number of days each patient will stay at the hospital next year.
If $\mu = 8$ and $\sigma = 2$, then the proportion of patients who will stay less than 7 days is equal to 0.3085.

**1.10.9** NH: $\mu_Y \leq 540$ *vs* AH: $\mu_Y > 540$

**1.10.10** $t_C = 19.1663$

Reject NH.

**1.10.11** The $P$-value is $< 0.0005$.

**1.10.12** $C[594.6 \leq \mu_Y] = 0.95$

**1.10.13** **a** Not defined. **b** Defined. **c** Defined. **d** Defined.

**1.10.14** **a** Cannot be done.

**b** $C - B^T = \begin{bmatrix} 10 & 0 & 0 \\ -1 & -7 & 4 \end{bmatrix}$

**c** $AB = \begin{bmatrix} 16 & 0 \\ 16 & 9 \\ 17 & 8 \end{bmatrix}$

**d** $CA^T = \begin{bmatrix} 76 & 46 & 37 \\ 1 & 19 & 35 \end{bmatrix}$

**1.10.15** $K = \begin{bmatrix} 11237 & 212.08 \\ 212.08 & 4.34 \end{bmatrix}$

**1.10.16** $K^{-1} = \begin{bmatrix} 0.00114 & -0.05572 \\ -0.05572 & 2.95214 \end{bmatrix}$

**1.10.18** $\mu_Y(3, -1) = 39$

**1.10.19** $\beta_0 + \beta_1 + 3\beta_2 - 2\beta_3$

**1.10.20** **a** linear in $\beta_0$ and $\beta_1$.

**b** linear in $\beta_1$ and $\beta_2$ but not linear in $\beta_0$.

**c** linear in $\beta_0$ and $\beta_2$ but not linear in $\beta_1$.

**d** linear in all the $\beta_i$, $i = 0, 1, 2, 3$, simultaneously.

**2.2.4** $P_Y(490) = 2.6956$, $P_Y(625) = 3.349$

**2.2.5** 0.242

**2.2.6** $x \geq 2.676/.00484 = 552.892562$, say 553 rounded to the nearest integer.

**2.4.4** Eleven cars in the population were driven 14,300 miles the first year. Eleven cars in the population were driven 7,100 miles the first year.

**2.4.5** Yes, there are 17 cars that were driven 9,200 miles the first year.

**2.4.6** There are 22 cars that were driven 8,700 miles during the first year. The mean maintenance cost for these cars is $433.318, and the standard deviation is $29.2177.

**2.4.8** There are 18 cars in the subpopulation with $X_2 = 10{,}000$. The mean maintenance cost for this subpopulation is $464.222, and the standard deviation is $30.5024.

**2.4.11** **a** Simple random sampling. **b** Sampling with preselected $X$ values.

**3.2.6** The sample in Problem 3.2.1 is preferred because the sample size is larger.

**3.4.1** $\hat{\sigma} = 0.916696$, $\hat{\beta}_1 = 0.530253$, $\hat{\beta}_0 = -0.227365$

**3.4.2** (1) On the average, the crystals grow 0.530 gram per hour.
(2) If a crystal is allowed to grow for 15 hours, its predicted size (weight) is 7.72643 grams.
(3) The additional dollars that the 24-hour crystal is expected to fetch is $138.62.

**3.4.3** $\hat{\mu}_Y(19) = 9.847$ grams. $\hat{\mu}_Y(25) = 13.029$ grams. For 40 hours it is not safe to use the estimated regression function based on these data because the data do not include 40 hours or any number close to 40 hours. Thus an excessive amount of *extrapolation* would be needed and the results would not be reliable in general.

**3.4.4** **a** $\hat{\beta}_1 = 0.03291$, $\hat{\beta}_0 = 158.49$, $\hat{\mu}_Y(x) = 158.49 + 0.03291x$

**b** The buyer would be interested in $Y(13000)$ and not $\mu_Y(13000)$. $\hat{Y}(13000) = \$586.28$. This is also the estimate for $\mu_Y(13000)$.

**c** Here the quantity of interest is $\mu_Y(16000)$.

**3.4.5** $\bar{x} = 6.00$, $\bar{y} = 68.7308$, $SXY = 1494.0$, $SSX = 364$, $SSY = 6295.12$, $\hat{\beta}_0 = 44.10$, $\hat{\beta}_1 = 4.10$, $SSE = 163.1$, $\sigma$ is unknown but $\hat{\sigma} = 2.607$

**3.4.6** **a**

| Row | y* | x* |
|---|---|---|
| 1 | -6.90 | -12 |
| 2 | -5.99 | -10 |
| 3 | -3.11 | -8 |
| 4 | -1.86 | -6 |
| 5 | -1.81 | -4 |
| 6 | -0.11 | -2 |
| 7 | -1.71 | 0 |
| 8 | 1.70 | 2 |
| 9 | 2.42 | 4 |
| 10 | 4.38 | 6 |
| 11 | 4.38 | 8 |
| 12 | 4.73 | 10 |
| 13 | 5.95 | 12 |
| 14 | 8.10 | 14 |

**b** $SSY = 265.947$, $SSX = 910$, $SXY = 482.53$, $\hat{\beta}_1^* = 0.5302$, $\hat{\sigma}^* = 0.9167$

**3.5.1** Yes. A straight line regression seems reasonable.

**3.5.2** $\hat{\beta}_0 = 1.54$, $\hat{\beta}_1 = 1.82$, $\hat{\sigma} = 1.514$

**3.5.5** The plot is not inconsistent with Gaussian assumptions.

**3.6.2** **a** $\hat{\beta}_1 = 0.03131$. Required difference in Problem 3.6.1 is estimated to be \$156.55.

**b** $C[126 \le 5000\beta_1 \le 187] = 0.90$

**c** \$1,470.23

**d** $C[548.22 \le \mu_Y(12{,}500) \le 588.48] = 0.90$

**e** $C[38.59 \le \sigma \le 59.69] = 0.80$. The interval is too wide to make a decision.

**3.6.3** You have 90% confidence that the two confidence intervals

$$-0.019 \le \beta_0 \le 0.228 \text{ and } .958 \le \beta_1 \le 1.02$$

are simultaneously correct. An investigator must decide if these are close enough to 0 and 1, respectively, in the problem being studied.

**3.6.4** **a** Use the following confidence statement to help make a decision. $C[0.4407 \le \beta_1 \le 0.5662] = 0.90$.

**b** $C[5.16 \le Y(15) \le 9.95] = 0.95$

**c** $C[2.61 \le Y(10) \le 7.46 \quad \text{and} \quad 10.07 \le Y(25) \le 15.10] \ge 0.90$

**3.7.1** **e** NH: $\beta_1 = 0$ and AH: $\beta_1 \ne 0$.
The $P$-value is between 0.001 and 0; i.e., $0 < P < 0.001$. Therefore NH is rejected and we conclude (at $\alpha = 0.05$) that shelf life is related to storage temperature.

**f** Zero degrees Celsius is very much outside the data range of temperatures, and so you should not extrapolate the results to zero.

**g** $C[705.7982 \le \mu_Y(15) \le 733.1072] = 0.95$

**h** $C[-15.3951 \le \beta_1 \le -12.1117] = 0.95$

Hence conclude that shelf life is related to storage temperature.

**i** NH: $\mu_Y(13) \le 650$ versus AH: $\mu_Y(13) > 650$. We get $t_C = \frac{746.9595 - 650}{7.7122} = 12.5723$. The $P$-value is between 0 and 0.0005. Thus the NH will be rejected at $\alpha = 0.05$.

**j** $C[730.6 \le \mu_Y(13) \le 763.3] = 0.95$. We can conclude that $\mu_Y(13)$ is greater than 650.

**3.8.1** **a** Analysis of Variance

| SOURCE | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 1 | 91645.2 | 91645.2 | 315.37 | < 0.01 |
| Error | 16 | 4649.7 | 290.6 | | |
| Total | 17 | 96294.9 | | | |

**b** The *P*-value is smaller than 0.01.

**c** $t_C = -17.758$

The *P*-value is between 0 and 0.001.

**e** From (b) we can conclude that $\beta_1$ is not equal to zero (if we use $\alpha = 0.05$ or $\alpha = 0.01$ for instance).

**3.8.2** **a** Analysis of Variance

| SOURCE | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 1 | 9337.7 | 9337.7 | 1161.31 | 0.000 |
| Error | 22 | 176.9 | 8.0 | | |
| Total | 23 | 9514.6 | | | |

**b** $F_C = 1162.81$. The *P*-value is less than 0.001. The NH: $\beta_1 = 0$ would be rejected at any commonly used $\alpha$ level.

**c** $t_C = 34.1$. The *P*-value is less than 0.001 so reject NH for $\alpha > .001$.

**d** $t_C^2 = (34.1)^2 = 1161.31 = F_C$ except for rounding errors.

**e** For any usual $\alpha$ value, NH: $\beta_1 = 0$ would be rejected.

**f** $C[1.48 \leq \beta_1 \leq 1.74] = 0.99$

**3.9.2** **a** Based on this plot the regression of $Y$ on $X$ seems to be linear.

**b** $\hat{\rho}_{Y,X} = 0.99$

**i** No

**ii** No

**iii** Yes

**c** $\hat{\sigma}_Y = 20.34$, and $\hat{\sigma} = 2.836$. $\hat{\sigma}$ can be used to decide whether $X$ is an adequate predictor of $Y$.

**3.10.1** **a** $\hat{\beta}_0 = 15.7$, $\hat{\beta}_1 = 0.414$, and $C[43.94 \leq \mu_Y(72) \leq 47.07] = 0.85$.

**c** The investigator concludes that soil temperature is not an important factor relative to the objectives.

**d** NH: $5\beta_1 \leq 5$ and AH: $5\beta_1 > 5$, or NH: $\beta_1 \leq 1$ and AH: $\beta_1 > 1$. $t_C = -6.01$, $n = 14$, so $df = 12$. *P*-value is greater than 0.9995. NH cannot be rejected using $\alpha = 0.10$.

**e** $\hat{Y}(75) = 46.75$. No valid prediction interval is available for $Y(75)$.

**f** None available

**g** $\hat{\sigma}_{V|U} = 3.649$

**3.10.2** **a** $\hat{\mu}_Y(u) = \hat{\beta}_0^* + \hat{\beta}_1^* u = 276 + 0.0259u$

**b** This is $\mu_Y(x) = \beta_0 + \beta_1 x$ with $x = 12{,}500$. No valid point or interval estimate of $\mu_Y(x)$ is available since no valid point estimate of $\beta_0$ or $\beta_1$ is available.

**c** $\mu_Y(u) = \beta_0^* + \beta_1^* u$ with $u = 12{,}500$. $\hat{\mu}_Y(12{,}500) = \hat{\beta}_0^* + \hat{\beta}_1^*(12500) = \$599.60$

**d** No valid 90% confidence interval is available.

**e** A valid confidence statement is $C[\$567.60 \leq \mu_Y(12{,}500) \leq \$631.50] = 0.95$.

**3.11.1** **b** $\hat{\beta}_1 = 0.504$ and $\hat{\sigma} = 1.02$

**c** $\beta_1$

**d** $C[0.482 \leq \beta_1 \leq 0.525] = 0.80$

**e** $C[0.78 \leq \sigma \leq 1.52] = 0.90$

**f** The investigator would conclude that $\mu_Y(x)$ is an adequate prediction function for this problem.

**3.11.2** The investigator would conclude that the chemical analysis provides unbiased estimates for As.

**3.12.2** **a** If assumptions (B) are valid, then all of the parameters listed have valid estimates.

**b**

$$\hat{\beta}_0 = 1.20260$$

$$\hat{\beta}_1 = 0.00257942$$

$$\hat{\sigma} = 0.298319$$

$$\hat{\mu}_Y(x) = 1.20260 + 0.00257942x$$

$$\hat{\mu}_Y = \bar{y} = 2.523$$

$$\hat{\mu}_X = \bar{x} = 511.90$$

$$\hat{\sigma}_X = 169.2$$

$$\hat{\sigma}_Y = 0.519$$

$$\hat{\rho}_{Y,X} = 0.841$$

**c**

$$\sigma: \sqrt{\frac{SSE}{15.507}}, \sqrt{\frac{SSE}{2.733}}, \quad \textit{i.e.}, (0.214, 0.510)$$

$$\sigma_Y: \sqrt{\frac{SSY}{16.919}}, \sqrt{\frac{SSY}{3.325}}, \quad \textit{i.e.}, (0.3787, 0.8543)$$

$$\sigma_X: \sqrt{\frac{SSX}{16.919}}, \sqrt{\frac{SSX}{3.325}}, \quad \textit{i.e.}, (123.4, 278.4)$$

**f** $50\beta_1$

**g** $C[0.06122 \le 50\beta_1 \le 0.19672] = 0.95$

**h** $C(0.12244 \le 100\beta_1 \le 0.39344] = 0.95$

**3.12.3** **b** The regression equation is

$$\hat{\mu}_Y(x) = 90.71 + 1.4272x$$

**d** $SE(\hat{\beta}_0) = 17.42$ and $SE(\beta_1) = 0.1777$

**e** $\hat{\sigma} = 29.56$

**f** $C[\beta_0 \le 120.916] = 0.95$

**g** $C[0.9158 \le \beta_1 \le 1.9386] = 0.99$

No. It appears that $\beta_1$ is not zero because the confidence interval above excludes zero.

**h** $C[22.34 \le \sigma \le 43.72] = 0.95$

**k** An upper 97.5% confidence bound for $\mu_Y(25)$ is 154.58.

**l** $\hat{\rho}_{Y,X} = 0.884$. $C[0.8 \le \rho_{Y,X} \le 0.95] = 0.95$. We conclude that the regression function $\mu_Y(x) = \beta_0 + \beta_1 x$, using fat intake as the predictor variable, is better than using $\mu_Y$ for predicting cholesterol.

**m** $t_C = -3.2234$. The *P*-value is greater than 0.995. The null hypothesis cannot be rejected using an $\alpha$ value of 0.01.

**n** **Accept** the hypothesis that $\beta_1$ is less than 2!

**3.12.4** **a** $\hat{\sigma}$ is 29.56 and $\hat{\sigma}_Y$ is 61.6.

**b** The analysis of variance table is

```
Analysis of Variance
SOURCE        DF        SS        MS        F        P
Regression     1     56368     56368    64.50    0.000
Error         18     15731       874
Total         19     72099
```

**c** $F_C = 64.50$

**d** The $P$-value is less than 0.01.

**e** $C[1.67 \leq \sigma_Y/\sigma \leq 3.20] = 0.95$.

**3.12.5** **a** NH: $\beta_1 \leq 0.8$ AH: $\beta_1 > 0.8$.

**b** $t_C = 3.53$, $P$-value $< 0.005$.

**c** The manufacturer's claim seems to be incorrect.

**d** $C[45.8 \leq 50\beta_1 \leq 96.9] = 0.99$

**4.2.1** **b** $E_I = 0.300$ for population item $I = 962$. $E_I = 0.900$ for population item $I = 1376$.

**d** $\sigma_Y(240{,}18) = 1.7076$

**e** $\sigma_E(240{,}18) = 1.7076$ including items 962 and 1376 in the calculations.

**f** Yes.

**4.2.2** 41.0

**4.2.4** $\mu_Y(280{,}18) = 33.0$

**4.4.2** $X^TX = \begin{bmatrix} 10 & 2410 & 138 \\ 2410 & 587900 & 32860 \\ 138 & 32860 & 2020 \end{bmatrix} \quad X^TY = \begin{bmatrix} 297.7 \\ 73497.0 \\ 3910.6 \end{bmatrix}$

**4.4.3** $(X^TX)^{-1} = \begin{bmatrix} 16.3005 & -0.0504 & -0.2930 \\ -0.0504 & 0.0002 & 0.0006 \\ -0.2930 & 0.0006 & 0.0107 \end{bmatrix} \quad \hat{\beta} = \begin{bmatrix} -0.64546 \\ 0.18721 \\ -1.06533 \end{bmatrix}$

**4.4.5** $\hat{\mu}_Y = 29.77$, $\hat{\mu}_{X_1} = 241$, $\hat{\mu}_{X_2} = 13.80$
$\hat{\sigma}_Y = 7.80$, $\hat{\sigma}_{X_1} = 28.07$, $\hat{\sigma}_{X_2} = 3.58$

**4.4.6** $\hat{\sigma} = 1.117$

**4.4.8** $\mu_Y(280{,}19)$

**4.4.10** $\hat{\mu}_Y(280{,}18) = 32.595$

**4.4.12** About 4.55% of the plastic containers have a strength greater than 31 if they are manufactured with a temperature of 250° and a pressure of 16.

**4.5.2** Based on the plots we conclude the data are *not inconsistent* with the assumption that $\{(Y, X_1, X_2, X_3, X_4)\}$ is Gaussian.

**4.5.3** There is no reason to seriously doubt that the set of standardized residuals is a simple random sample from a standard Gaussian population.

**4.6.2** $\hat{\mu}_{X_1} = 4033.2$, $\hat{\mu}_{X_2} = 2.5588$, $\hat{\mu}_{X_3} = 1889.1$, $\hat{\mu}_Y = 619.41$.
$\hat{\sigma}_{X_1} = 1082.2$, $\hat{\sigma}_{X_2} = 1.5013$, $\hat{\sigma}_{X_3} = 692.68$, $\hat{\sigma}_Y = 334.9$[illegible].

**4.6.3** $\hat{\beta}_0 = -358.4$, $\hat{\beta}_1 = 0.0751$, $\hat{\beta}_2 = 55.09$, $\hat{\beta}_3 = 0.2811$.
$SE(\hat{\beta}_0) = 198.7$, $SE(\hat{\beta}_1) = 0.1361$, $SE(\hat{\beta}_2) = 29.05$, $SE(\hat{\beta}_3) = 0.2261$.
$\hat{\mu}_Y(x_1, x_2, x_3) = -358.4 + 0.0751x_1 + 55.09x_2 + 0.2811x_3$

**4.6.5** $\hat{\sigma}_Y = 334.92$

**4.6.7** $C[\mu_Y(4000, x_2, x_3) - \mu_Y(3000, x_2, x_3) \le \$306.12] = 0.95$

**4.6.8** $C[\mu_Y(x_1, x_2, 2000) - \mu_Y(x_1, x_2, 1500) \le \$332.43] = 0.95$

**4.6.9** $\hat{Y}(3200, 6, 2800) = \$999.60$

**4.6.10** $C[Y(3200, 6, 2800) \le 1361.46] = 0.90$

**4.6.11** $C[\mu_Y(3200, 6, 2800) \le 1314.89] = 0.90$

**4.6.13** $\hat{\sigma} = 135.42$ and $C[108.21 \le \sigma \le 181.01] = 0.95$

**4.7.1** **b** NH: $\beta_1 = 0$ AH: $\beta_1 \ne 0$. The $P$-value is 0.585.

**4.7.3** $X_C^2 = 550163/(50)^2 = 220.07$. The $P$-value is greater than 0.99.

**4.8.1** $P$-value $< 0.01$.

**4.8.2** $P$-value $< 0.01$.

**4.8.3** $P$-value $< 0.01$.

**4.9.1** **c** $\hat{\rho}^2_{Y(X_1,X_2,X_3)} = 0.905$
$\hat{\rho}^2_{Y(X_4,X_5,X_6,X_7)|X_1,X_2,X_3} = 0.127$
$\hat{\rho}^2_{Y(X_1,X_2,\ldots,X_7)} = 0.917$

**d** $\dfrac{\hat{\sigma}_B}{\hat{\sigma}_A} = 0.927$

**4.10.1** **a** **i** $C[0.255 \le \sigma_{Y|X_1} \le 0.498] = 0.95$, $C[0.347 \le \sigma_{Y|X_3} \le 0.679] = 0.95$

**ii** $C[0.376 \le \sigma_{Y|X_1}/\sigma_{Y|X_3} \le 1.435] = 0.90$

**b** **i** $C[0.365 \le \sigma_{Y|X_2} \le 0.715] = 0.95$, $C[0.384 \le \sigma_{Y|X4} \le 0.751] = 0.95$

**ii** $C[0.486 \le \sigma_{Y|X_2}/\sigma_{Y|X_4} \le 1.862] = 0.90$

**4.11.1** **a** $m = 4$

**b** $n = 24$, $n_1 = 6$, $n_2 = 6$, $n_3 = 6$, and $n_4 = 6$

**c** $\hat{\mu}_1 = 11.567$, $\hat{\mu}_2 = 10.033$, $\hat{\mu}_3 = 8.217$, $\hat{\mu}_4 = 6.450$
$\hat{\sigma}_1 = 2.342$, $\hat{\sigma}_2 = 2.655$, $\hat{\sigma}_3 = 2.070$, $\hat{\sigma}_4 = 2.145$

**e** $\hat{\theta}_1 = -0.075$, $\hat{\theta}_2 = 0.1083$, $\hat{\theta}_3 = 0.0083$, and $\hat{\theta}_4 = -0.0417$

**h** 1.91 units

**i** Yes, for the $X$ values studied.

**j** $F_C = 0.011$. The $P$-value is greater than 0.20.

**4.12.1** **c** $\hat{\mu}^{(t)}_{log_{10}(C)}(30) = 0.0404$

**d** For this problem a quadratic term cannot be ignored.

**e** NH: $\beta_2 = 0$ against NH: $\beta_2 \ne 0$. The $P$-value is less than 0.001.

**4.12.2** **b** **i** $\hat{\beta}_0^{(A)} = 200.68$, $\hat{\beta}_1^{(A)} = 0.014623$, $\hat{\sigma}_{Y|X} = 22.29$.
$\mu_Y^{(A)}(x) = \hat{Y}(x) = 200.68 + 0.014623x$.

**ii** $SS$ (pure error) = 1804.436, $df$ (pure error) = 8, $MS$ (pure error) = 225.55.

**v** Yes

**vi** NH: The regression function is $\mu_Y(x) = \beta_0 + \beta_1 x$ against AH: The regression function is not $\mu_Y(x) = \beta_0 + \beta_1 x$. The $P$-value is between 0.025 and 0.05.

**c** $\hat{\beta}_0 = 258.72$, $\hat{\beta}_1 = 0.003309$, $\hat{\beta}_2 = 0.00000042$, $\hat{\sigma}_{Y|X_1,X_2} = 12.85$

**4.12.3** **c** By preselecting $X_1$ and $X_2$ values.

**e** $\hat{\beta}_0 = 1.697$, $\hat{\beta}_1 = -0.02030$, $\hat{\beta}_2 = -0.5477$, $\hat{\beta}_3 = 0.0151$. $\hat{\sigma}_{Y|X_1,X_2,X_3} = 0.3724$
$\hat{\mu}_Y(x_1, x_2, x_3) = \hat{Y}(x_1, x_2, x_3) = 1.697 - 0.0203x_1 - 0.5477x_2 + 0.0151x_3$

**f** NH: $\beta_2 = \beta_3 = 0$. AH: at least one of $\beta_2$, $\beta_3$ is not zero. The $P$-value is less than 0.01.

**g** $\hat{\mu}_Y(65, 10, 650) = 4.72$ and $C[4.44 \le \mu_Y(65, 10, 650) \le 4.99] = 0.95$

**4.12.4** **b** **ii** $\hat{\sigma}_A = 1.004$

**iii**

| $\beta_i$ | Lower | Upper |
|---|---|---|
| $\beta_0$ | −137.014 | −19.5212 |
| $\beta_1$ | 0.237369 | 2.50626 |
| $\beta_2$ | 0.348319 | 1.21653 |
| $\beta_3$ | 0.7555 | 1.34732 |
| $\beta_4$ | −0.494088 | 0.254260 |
| $\beta_5$ | −0.192064 | 0.374936 |
| $\beta_6$ | −0.263157 | 0.439846 |
| $\beta_7$ | −0.439239 | 0.23575 |

**v** $C[2.106 \le \sigma_Y \le 4.045] = 0.95$, $C[0.720 \le \sigma_A \le 1.657] = 0.95$, and $C[1.271 \le \sigma_Y/\sigma_A \le 5.618] \ge 0.90$

**vi** 0.917

**vii** 69.347 inches

**ix** $C[65.98 \le \mu_Y(20, 60, 72, 61, 71, 62, 70)] = 0.95$

**x** $C[66.5 \le \mu_Y(20, 60, 72, 61, 71, 62, 70)] = 0.95$

**xi** Standard errors are not the same.

**c** **i** $\hat{\beta}_0^C = -61.20$, $\hat{\beta}_2^C = 0.8947$, $\hat{\beta}_3^C = 1.0556$, $\hat{\sigma}_{Y|X_2,X_3} = 1.131$.

**ii** $\hat{\beta}_0^D = -78.23$, $\hat{\beta}_1^D = 1.3503$, $\hat{\beta}_2^D = 0.6925$, $\hat{\beta}_3^D = 1.1025$. $\hat{\sigma}_{Y|X_1,X_2,X_3} = 0.9305$.

**iv** $\mu_Y^{(C)}(x_2 + 1, x_3) - \mu_Y^{(C)}(x_2, x_3) = \beta_2^C$. A 95% lower confidence bound for $\beta_2^C$ is 0.587.

**vii** $\mu_Y^{(C)}(x_2, x_3 + 1) - \mu_Y^{(C)}(x_2, x_3) = \beta_3$. A 95% upper confidence bound for $\beta_3^C$ is 1.26.

**viii** NH: $\beta_3^C > 1$. AH: $\beta_3^C \le 1$. The $P$-value is greater than 0.20.

**x** $\hat{\rho}^2_{Y(X_3)|X_2} = 0.84$

**xi** $C[0.888 \le \sigma_{Y|X_2,X_3} \le 1.583] = 0.90$

**xii** $\hat{Y}(60, 72) = 68.5$, $C[65.5 \le Y(60,72) \le 71.4] = 0.95$

**xiii** $\hat{\mu}_Y^{(C)}(60, 72) = 68.5$, $C[67.04 \le \mu_Y^{(C)}(60, 72) \le 69.93] = 0.90$

**d** **ii** $C[67.1 \le Y^{(C)}(20, 60, 72) \le 72.3] = 0.95$

**iii** $C[68.0 \le \mu_Y^{(C)}(20, 60, 72) \le 71.4] = 0.95$

**iv** $\hat{\rho}^2_{Y(X_1)|X_2,X_3} = 0.3627$

**5.2.1** **a** A plot of $Y$ against $X$ reveals no outliers.

**b** Observation 12 is a candidate for a possible outlier and should be examined.

**5.2.2** For observation 23, the studentized deleted residual is $T_{23} = 21.6335$ and the standardized residual is $r_i = 5.5577$.

**5.3.1** **a** $n = 36$. $p = 3$. $2p/n = 0.166667$. **b** Observation 13 has a hat value of 0.191368. Observation 21 has a hat value of 0.200206.

**5.3.3** No change. No change.

**5.4.1** **a** $h_{3,3,} = 0.14463$
**b** $\text{DFFITS}_4 = 0.94814$
**c** $c_4 = 0.26849$
**d** $\hat{Y}(3.5) = 17.4884$
**e** $\hat{r}_3 = -0.64454$
**f** $\hat{e}_3 = -1.08058$
**g** $T_6 = -0.60789$

**5.4.2** **a** 2.52849. Yes
**b** 0.81405. Yes
**c** 0.26849. No
**d** 0.94814. Yes

**5.4.3** **a** $\hat{Y}(13) = 21.562$
**b** $\hat{Y}_{(-6)}(13) = 21.76$
**c** $\hat{\sigma} = 3.476$
**d** $\hat{\sigma}_{(-6)} = 1.967$

**5.4.4** $T_6 = -3.08081$ in Exhibit 5.4.5.

**5.4.5** $h_{6,6} = 0.97319$ in (5.3.3) and in Exhibit 5.4.5.

**5.6.1** **a** No.
**b** Sample items 12 and 16 should be examined further as possible outliers.
**c** Items 5, 12, 16, and 17 have DFFITS larger than $2\sqrt{p/n} = 1.26$ and should be examined as influential observations. Cook's distance for item 12 is greater than $F_{.5:8,12} = 0.97$. Hence items 5, 12, 16, and 17 should be examined.
**e** No

**6.2.1** **a** $\hat{Y}(12.0, 22, 9.3) = \$90.58$
**b** $\hat{\mu}_Y(12.0, 22, 9.3) = \$90.58$
**c** $C[\$0.00 \le Y(12.0, 22, 9.3) \le \$200.93] = 0.95$
**d** $C[\$46.46 \le \mu_Y(12.0, 22, 9.3) \le \$134.70] = 0.95$

**6.2.3** \$355.74

**6.2.4** **a** $\hat{Y}(15) = 7.553$ grams
**b** $C[5.59 \le Y(15) \le 9.51] = 0.90$
**c** $\hat{Y}(3) = 1.512$, $\hat{Y}(5) = 2.519$, and $\hat{Y}(13) = 6.546$, so $\hat{Y}_S = 10.577$.
**d** $C[6.66 \le Y_S] = 0.95$
**g** $\hat{\mu}_Y(20) = 10.07$ grams
**h** $C[9.47 \le \mu_Y(20) \le 10.67] = 0.90$

**6.2.5** **a** $\mu_Y(30)$, $\mu_Y(60)$
**b** $\hat{\mu}_Y(60) = 9.92$
**c** $C[9.33 \le \mu_Y(60) \le 10.51] = 0.90$

**6.3.1** **a** $\hat{\lambda}_{0.85}(3.0) = 0.868$
**b** We have 90% confidence that at least a proportion $p = 0.85$ of these subpopulation $Y$ values is between 0.392 and 1.574.
**c** We have 90% confidence that $\lambda_{0.15}(3.0)$ is between $-2.024$ and $-0.843$.

**6.3.2** We have confidence of at least 90% that a proportion 0.7 of the subpopulation values at $X = 3.0$ are between $L$ and $U$ where $L = -2.024$ and $U = 1.574$.

**6.3.3** **b** No. It is better to compare this baby's weight with $\lambda_p(30)$, for some suitable percentile $p$.

**c** iv since $k = \lambda_{0.10}(30)$.

**d** $\hat{\lambda}_{0.10}(30) = 5.67$ pounds and $C[4.454 \le \lambda_{0.10}(30) \le 6.604] = 0.95$.

**e** $\lambda_p(x), x = 60$ .

**f** $p = 0.05$

**g** We have confidence of 95% that at most 5% of babies who are 60 days old weigh less than 5.45 pounds.

**6.4.1** $\hat{x}_0 = 66.5$ degrees

**6.4.2** $C[65.07 \le x_0 \le 68.03] = 0.99$

**6.4.3** The estimate of her actual temperature is 100.1 degrees.

**6.4.4** $C[99.63 \le x_0 \le 100.59] = 0.90$

**6.4.5** $\hat{x}_0 = 9.93$

**6.4.6** $C[8.65 \le x_0 \le 11.05] = 0.90$

**6.4.7** $\hat{X}(y_0) = \hat{X}(10.5) = 4.718$. You should consider the model $\mu_X(y) = \beta_0^* + \beta_1^* y$ since assumptions (B) apply.

**6.4.8** $C[3.340 \le X(10.5) \le 6.096] = 0.90$

**6.5.1** **a** $\hat{\alpha}_1 = -1.882, \hat{\beta}_1 = 1.098$
$\hat{\alpha}_2 = -1.871, \hat{\beta}_2 = 1.106$
$\hat{\alpha}_3 = -2.472, \hat{\beta}_3 = 1.199$

**b** $\hat{\sigma}_1 = 0.5536, \hat{\sigma}_2 = 0.1766, \hat{\sigma}_3 = 0.3663$

**c** $\hat{\sigma} = 0.3966$

**d** $d^T = [0 \quad 1 \quad 0 \quad -1 \quad 0 \quad 0]$

**e** $d^T = [0 \quad 0 \quad 1 \quad 0 \quad -1 \quad 0]$

**f** $m = 2$

**g** $m = 3$

**h** $m = 2$

**i** One has at least 90% confidence that the following are correct.

$$-0.9554 \le \alpha_1 - \alpha_2 \le 0.9340$$
$$-0.1220 \le \beta_1 - \beta_2 \le 0.1061$$

**6.5.2** The confidence intervals are

$$-1.18876 \le \alpha_1 - \alpha_2 \le 1.16731$$
$$-0.54474 \le \alpha_1 - \alpha_3 \le 1.72542$$
$$-0.40567 \le \alpha_2 - \alpha_3 \le 1.60779$$
$$-0.15022 \le \beta_1 - \beta_2 \le 0.13428$$
$$-0.23798 \le \beta_1 - \beta_3 \le 0.03659$$
$$-0.16820 \le \beta_2 - \beta_3 \le -0.01726$$

The rates of growth are "equivalent" for this problem. Results for intercepts are not conclusive and a larger sample size is need.

**6.5.3** $\mu_Y^{(1)}(0) - \mu_Y^{(2)}(1.5) = \alpha_1 - \alpha_2 - 1.5\beta_2$

**6.5.4** $m = 30$

**6.5.5** *The investigator would like to know whether the following inequalities are true.*

$$|\mu_Y^{(1)}(1) - \mu_Y^{(2)}(1)| \le 2, \quad |\mu_Y^{(1)}(30) - \mu_Y^{(2)}(30)| \le 2$$
$$|\mu_Y^{(1)}(1) - \mu_Y^{(3)}(1)| \le 2, \quad |\mu_Y^{(1)}(30) - \mu_Y^{(3)}(30)| \le 2$$
$$|\mu_Y^{(2)}(1) - \mu_Y^{(3)}(1)| \le 2, \quad |\mu_Y^{(2)}(30) - \mu_Y^{(3)}(30)| \le 2$$

**6.6.1** $\hat{x}_0 = 0.256$

**6.6.2** $\hat{x}_0 = -0.3387$

**6.6.3** $C[-1.359 \le x_0 \le 1.414] = 0.95$

**6.6.4** $\hat{x}_0 = -1.346$

**6.6.5** The confidence region is $-\infty$ to $+\infty$.

**6.7.1** **b** $\hat{\beta}_0 = 2.1628,\ \hat{\beta}_1 = 0.32785,\ \hat{\beta}_2 = -0.005151,\ \hat{\sigma} = 0.6300$

**c** $\hat{x}_0 = 31.8$

**d** $C[29.9 \le x_0 \le 34.0] = 0.90$

**6.8.1**

$$X = \begin{bmatrix} 1 & 100 & 0 \\ 1 & 125 & 0 \\ 1 & 150 & 0 \\ 1 & 175 & 0 \\ 1 & 200 & 0 \\ 1 & 225 & 0 \\ 1 & 250 & 0 \\ 1 & 270 & 5 \\ 1 & 270 & 30 \\ 1 & 270 & 55 \\ 1 & 270 & 80 \\ 1 & 270 & 105 \\ 1 & 270 & 130 \end{bmatrix}$$

**6.8.2** **a** $\hat{\sigma} = 0.2795$. Also $C[0.1953 \le \sigma \le 0.4905] = 0.95$.

**b** $\hat{\alpha}_1 = 0.780,\quad C[0.187 \le \alpha_1 \le 1.37] = 0.90$
$\hat{\beta}_1 = 0.0120,\quad C[0.0091 \le \beta_1 \le 0.0150] = 0.90$
$\hat{\alpha}_2 = 6.823,\quad C[5.531 \le \alpha_2 \le 8.115] = 0.90$
$\hat{\beta}_2 = -0.0103,\quad C[-0.0143 \le \beta_2 \le -0.00640] = 0.90$

**c** $\hat{\mu}_Y(175) = 2.89,\quad C[2.6641 \le \mu_Y(175) \le 3.1071] = 0.95$
$\hat{\mu}_Y(375) = 2.94,\quad C[2.606 \le \mu_Y(375) \le 3.279] = 0.95$

**6.8.4** Neither model appears to be inconsistent with these data.

**6.9.1** **a** All of these parameters can be validly estimated.

**b** $\hat{\mu}_Y = 50.472,\ \hat{\mu}_X = 51.492.\ \hat{\sigma}_Y = 0.882,\ \hat{\sigma}_X = 0.779,\ \hat{\sigma} = 0.3949.\ \hat{\beta}_0 = -2.226,$
$\hat{\beta}_1 = 1.0234$

**c** $\hat{Y}(52.08) = 51.074$. We have 95% confidence that $Y(52.08) \ge 50.322$.

**d** $\hat{Y}_S = 203.896$ seconds. We have 99% confidence that $Y_S \ge 201.320$.

**e** $C[201.810 \le Y_S \le 205.982] = 0.95$

**6.9.2** **a** $\hat{\lambda}_{0.99}(210) = 149.4.\ C[145.2 \le \lambda_{0.99}(210) \le 156.9] = 0.95$

**c** $C[124.14 \le \lambda_{0.99}(160) \le 136.12] = 0.95$

**6.9.3** **a** $\hat{\mu}_Y(x) = 101 - 3.04x$

**b** $\hat{x}_0 = 15.08,\ C[10.8 \le x_0 \le 19.6] = 0.95$

**6.9.4** **b** $\hat{\mu}_Y(x) = 51.9 - 0.27x$

**c** $\hat{x}_0 = 62.20.\ C[59.59 \le x_0 \le 65.17] = 0.95$

**d** $C[39.85 \le \lambda_{0.01}(25) \le 42.45] = 0.95$

**6.9.5** **a** $\hat{\mu}_Y^{(1)}(x) = 23.5 + 0.0643x$, $\hat{\mu}_Y^{(2)}(x) = 24.5 + 0.0345x$, $\hat{\mu}_Y^{(3)}(x) = 29.3 + 0.0278x$.
$\hat{\sigma}_1 = 2.823$, $\hat{\sigma}_2 = 3.171$, $\hat{\sigma}_3 = 4.089$

**b** $\hat{\sigma} = 3.476$

**d** $\mu_Y^{(1)}(1000) - \mu_Y^{(2)}(1000)$
$\mu_Y^{(1)}(1000) - \mu_Y^{(3)}(1000)$
$\mu_Y^{(2)}(1000) - \mu_Y^{(3)}(1000)$

**6.9.6** **b** $\mu_Y^{(1)}(x) - \mu_Y^{(2)}(x) = \alpha_1 - \alpha_2 + \beta_1 x - \beta_2 x$

**c** $\mu_Y^{(1)}(0) - \mu_Y^{(2)}(0) = \alpha_1 - \alpha_2$. $C[-2.46 \le \alpha_1 - \alpha_2 \le 1.81] = 0.95$

**d** $\beta_1 - \beta_2$. $C[-0.476 \le \beta_1 - \beta_2 \le -0.178] = 0.95$

**6.9.7** **a** $\mu_Y^{(1)}(25) - \mu_Y^{(2)}(25)$

**b** $\mu_Y^{(1)}(60) - \mu_Y^{(2)}(60)$

**c** $\beta_1 - \beta_2$

**d** $\hat{\mu}_Y^{(1)}(25) - \hat{\mu}_Y^{(2)}(25) = 2.515$, $\hat{\mu}_Y^{(1)}(60) - \hat{\mu}_Y^{(2)}(60) = -0.386$, $\hat{\beta}_1 - \hat{\beta}_2 = -0.083$

**f** $\hat{x}_0 = 55.3448$ and $C[49.88 \le x_0 \le 65.20] = 0.90$

**6.9.8** **a** *MSE* (cubic) = 0.01092
*MSE* (quadratic) = 0.01194
*MSE* (linear) = 0.07059
With these mean squared errors, and after examining the plots, we choose the quadratic model.

**b** $\hat{\beta}_0 = 0.334$, $\hat{\beta}_1 = 0.0399$, $\hat{\beta}_2 = -0.000429$, $\hat{\sigma}_{quadratic} = \hat{\sigma} = 0.1093$

**c** $\mu_Y(12) = \beta_0 + 12\beta_1 + 144\beta_2$

**d** $\hat{\mu}_Y(12) = 0.7514$ and $C[0.696 \le \mu_Y(12) \le 0.807] = 0.80$

**e** $C[0.594 \le Y(12) \le 0.909] = 0.80$

**6.9.9** **a** By preselecting $X$ values.

**b** $\hat{\mu}_Y(x) = 0.085 + 0.379x - 0.156x^2$

**c** NH: $\beta_2 = 0$. AH: $\beta_2 \ne 0$

**d** The $P$-value is less than 0.001, so we reject NH at the 0.05 level.

**e** $x_0 = 1.218$

**f** $C[1.128 \le x_0 \le 1.366] = 0.95$

**g** $C[13.42 \le z_0 \le 23.24] = 0.95$

**6.9.10** **a** $\hat{\alpha}_1 = 7.08$, $\hat{\beta}_1 = 0.06732$, $\hat{\alpha}_2 = 9.581$, $\hat{\beta}_2 = 0.0423$

**b** $\beta_1 - \beta_2$

**c** $\hat{\beta}_1 - \hat{\beta}_2 = 0.025$, $C[0.0193 \le \beta_1 - \beta_2 \le 0.0307] = 0.90$

**d** $\hat{\beta}_0 = 6.92$, $\hat{\beta}_1 = 0.08$, $\hat{\beta}_2 = -0.00013$. *MSE* (spline) = 0.047, *MSE*(quadratic) = 0.050. The two models fit equally well according to the data. The residual plots indicate that neither model is inconsistent with assumptions (A).

**7.3.1** **a** $X_2$

**b** $X_2, X_5$

**c** $X_1, X_2, X_5$

**d** $X_1, X_2, X_3, X_4, X_5, X_6, X_7$
$X_1, X_2, X_3, X_5, X_6, X_7$
$X_1, X_2, X_4, X_5, X_6, X_7$

**e** $X_1, X_2, X_3, X_5, X_6, X_7$
$X_1, X_2, X_5, X_6, X_7$
$X_1, X_2, X_3, X_4, X_5, X_6, X_7$

**f** Same as (e).

**g** $X_1, X_2, X_5, X_6, X_7$
$X_1, X_2, X_3, X_5, X_6, X_7$
$X_1, X_2, X_3, X_4$

**h** $X_1, X_2, X_5, X_6, X_7$
$X_1, X_2, X_3, X_5, X_6, X_7$
$X_1, X_2, X_4, X_5, X_6, X_7$
$X_1, X_2, X_3, X_4, X_5, X_6, X_7$

**i** Variables $X_1, X_2, X_5, X_6, X_7$ occur in most of the better fitting models. The model with $X_1, X_2, X_5, X_6, X_7$ appears to be a good model based on the criteria $R^2$, $adj\text{-}R^2$, $s$, $C_p$, and $C_p - p$.

**7.4.1** **a** $X_1, X_2, X_3$

**b** $X_1, X_2$

**c** $X_1, X_2, X_3$

**d** $X_1, X_2$

**f** $X_1, X_2, X_3$

**g** $X_1, X_2, X_3$

**7.4.2** **a** None

**b** $X_1, X_2, X_3$

**c** $X_2, X_3$

**d** $X_2, X_3$

**h** *F-in* must be greater than or equal to *F-out*.

**i**

| *Step* | **Variables in Model** |
|---|---|
| 1 | $X_3$ |
| 2 | None |
| 3 | $X_1$ |

**7.4.4** **a** $X_1, X_2, X_3$

**b** $X_1, X_2, X_3$

**c** $X_1, X_2, X_3$

**d** $X_1, X_2, X_3$

**e**

| *Step* | **Variables in Model** |
|---|---|
| 1 | $X_1$ |
| 2 | None |
| 3 | $X_3$ |
| 4 | $X_2, X_3$ |
| 5 | $X_1, X_2, X_3$ |

**g**

| *Step* | **Variables in Model** |
|---|---|
| 1 | $X_1$ |
| 2 | None |
| 3 | $X_3$ |
| 4 | $X_2, X_3$ |
| 5 | $X_1, X_2, X_3$ |

**7.5.1** **a** $k = 5, m = 24, p = 3$

**b** $t_1 = 4, t_2 = 6, t_3 = 8, t_4 = 10, t_5 = 12$

**c**
$$X = \begin{bmatrix} 1 & 4 & 16 \\ 1 & 6 & 36 \\ 1 & 8 & 64 \\ 1 & 10 & 100 \\ 1 & 12 & 144 \end{bmatrix}$$

**d** $y_3 = [3.5 \quad 11.3 \quad 18.4 \quad 22.5 \quad 25.3]^T$
$y_{10} = [3.6 \quad 11.4 \quad 18.3 \quad 21.6 \quad 23.9]^T$

**e** $y_{15,j} = \alpha_{15} + \beta_{15}t_j + \gamma_{15}t_j^2 + e_{15,j}$ for $j = 1, 2, 3, 4, 5$.

**f** $C[-0.258 \leq \gamma \leq -0.234] = 0.90$

**g** $\hat{\mu}_Y(t) = -18.94 + 6.17t - 0.246t^2$

**h** $\hat{\mu}(8) = 14.710$

**i** $a^T = [0, 1, 0]$

**j** $a^T = [1, t, t^2]$

**k** $8\beta + 128\gamma$

**7.6.1** **a** If $s$ is used as the criterion, then the short list of the three best subset models is $(X_1, X_2, X_3, X_4)$; $(X_1, X_2, X_3)$; $(X_1, X_3, X_4)$.
If $R^2$ is used as the criterion, then the short list of three best subset models is $(X_1, X_2, X_3, X_4)$; $(X_1, X_2, X_3)$; $(X_1, X_3, X_4)$.
If $adj\text{-}R^2$ is used as the criterion, then the short list of three best subset models is $(X_1, X_2, X_3, X_4)$; $(X_1, X_2, X_3)$; $(X_1, X_3, X_4)$.

**b** **i**

| $C_p$ | Variables in Model |
|---|---|
| 30.5 | $X_1$ |
| 37.3 | $X_3$ |
| 41.7 | $X_2$ |
| 43.3 | $X_4$ |
| 5.6 | $X_1, X_3$ |
| 12.8 | $X_1, X_2$ |
| 17.3 | $X_1, X_4$ |
| 23.2 | $X_3, X_4$ |
| 24.5 | $X_2, X_3$ |
| 25.9 | $X_2, X_4$ |
| 4.2 | $X_1, X_2, X_3$ |
| 5.3 | $X_1, X_3, X_4$ |
| 10.4 | $X_1, X_2, X_4$ |
| 18.8 | $X_2, X_3, X_4$ |
| 5.0 | $X_1, X_2, X_3, X_4$ |

**iii** $X_1, X_2, X_3$

**iv** $(X_1, X_2, X_3, X_4)$; $(X_1, X_3, X_4)$; $(X_1, X_3)$

**c**

| *Step* | Variables in Model |
|---|---|
| 1 | 1 |
| 2 | 1, 3 |
| 3 | 1, 2, 3 |

**f** Variables $X_1$ and $X_3$ appear in all well-fitting models. The model with $X_1, X_2, X_3$ is a good final candidate.

**7.6.2** **a** $m = 20, k = 4, p = 3$

**b**

$$X = \begin{bmatrix} 1 & 8.0 & 64.00 \\ 1 & 8.5 & 72.25 \\ 1 & 9.0 & 81.00 \\ 1 & 9.5 & 90.25 \end{bmatrix}$$

**c** $a^T = [1, 0, 0]$

**d** $\mu_Y(t) = 27.87 + 3.39t - 0.092t^2$

**e** $C[-3.864 \leq \beta \leq 12.751] = 0.95$; $C[-0.534 \leq \gamma \leq 0.351] = 0.95$

**f** $\mu_Y(t) = \alpha + \beta t + \gamma t^2$

**g** $\mu_Y(8.5) = \alpha + 8.5\beta + 72.25\gamma$

**h** No. Need more data.

**8.2.1** **a** 0.0651 and 0.2500, respectively.

**c** $\hat{\beta}_0^{(w)} = 1.72$; $\hat{\beta}_1^{(w)} = 1.78$

**d** $\hat{\beta}_0 = 1.54$; $\hat{\beta}_1 = 1.82$

**8.2.2** $\hat{\mu}_Y^{(w)}(3.0) = 7.0499$; $\hat{\sigma}_0 = 0.334$; $\hat{\sigma}_Y(3.0) = 1.002$

**8.2.3** $2.5\beta_1$

**8.2.4** $C[3.125 \leq 2.5\beta_1 \leq 5.756] = 0.95$

**8.2.6** $C[0.252 \leq \sigma_0 \leq 0.506] = 0.90$

**8.2.7** $C[1.008 \leq \sigma_Y(4.00) \leq 2.024] = 0.90$

**8.3.1** **a** $n = 14$

**b** $m = 7$

**c** $w_2 = 2.24$; $v_2 = 21.8000$; $v_6 = 35.552$. $t_2 = 2.792$

**d** $q_2^* = 1.24643$; $q_4^* = 0.36414$; $q_2 = 0.36414$; $q_4 = 1.24643$

**e** $\hat{\beta}_0 = 1.24643$

**f** $\hat{\beta}_1 = 1.71875$

**g** $C[0.364 \leq \beta_0 \leq 3.146] = 0.88$

**h** $C[1.075 \leq \beta_1 \leq 2.366] = 0.88$

**i** By Theil's method, $\hat{\beta}_0 = 1.24643$ and $\hat{\beta}_1 = 1.71875$. By weighted least squares, $\hat{\beta}_0 = 1.72$ and $\hat{\beta}_1 = 1.78$.

**j** $\hat{\mu}_Y(5.0) = 10.3087$

**k** $\mu_Y(5.0) - \mu_Y(2.5) = 2.5\beta_1$

**l** $2.5\hat{\beta}_1 = 4.297$; $C[2.69 \leq 2.5\beta_1 \leq 5.92] = 0.88$

**8.3.2** Yes. A Gaussian population is symmetric and continuous.

**8.3.3** Yes

**8.3.4** Yes, but not necessarily

**8.3.5** Yes

**8.3.6** Yes, if $g(x) = 1$.

**8.3.7** Yes, if $g(x) = 1$.

**8.3.8** Yes

**8.3.9** Yes

**8.4.1** **a** Ordinary least squares (OLS) estimate is $\hat{\mu}_Y(x) = 1.99 + 0.912x$.

**b** $\hat{\mu}_Y^{(w)}(x) = 2.07 + 0.858x$

**c** $C[6.911 \le \mu_Y(6) \le 7.522] = 0.99$

**d** $C[6.0016 \le Y(6) \le 8.432] = 0.90$

**e** $C[0.0155 \le \sigma_0 \le 0.0254] = 0.80$
$C[0.556 \le \sigma_Y(6.0) \le 0.915] = 0.80$

**f** An estimate of $P[Y \ge 7]$ is 0.6245.

**8.4.2** **a** $\hat{\mu}_Y(x) = -6.132 + 1.082x$

**b** No, except student 8 has a large standardized residual = 3.46.

**d** **i** $-3.208 + 1.0099x$

**ii** $C[-15.06 \le \beta_0 \le 11.8] = 0.85$
$C[0.733 \le \beta_1 \le 1.121] = 0.85$

**iii** $\hat{Y}(75) = \hat{\mu}_Y(75) = 73.74$

**iv** $C[70.875 \le \mu_Y(75) \le 77.24] = 0.85$

**9.2.3** **a** This is a linear regression function.

**b** This is a linear regression function.

**c** This is a nonlinear regression function. It is linear in $\beta_0$ but is nonlinear in $\beta_1$.

**d** This is a linear regression function.

**e** This is a nonlinear regression function. It is simultaneously linear in $\beta_0$ and $\beta_1$, but it is nonlinear in $\beta_2$ and in $\beta_3$.

**9.3.1** $C[-0.3595 \le \beta_1 \le 0.4170] \approx 0.95$
$C[2.2470 \le \beta_2 \le 3.1996] \approx 0.95$
and
$C[0.3624 \le \beta_3 \le 1.0031] \approx 0.95$

**9.3.2** **a** $\hat{\beta}_1 = 1.9483$
$\hat{\beta}_2 = -1.2699$
$\hat{\beta}_3 = 14.3631$
$\hat{\sigma} = 0.1025$

**c** 1.8929671

**d** $C[0.5182 \le \theta \le 2.9690] \ge 0.90$ (approximately).

**9.3.3** **a** $\hat{\beta}_1 = 0.8183215325$, $\hat{\beta}_2 = -0.6776058218$, $\hat{\beta}_3 = 0.1632703464$, and $\hat{\sigma} = 0.3696$.

**c** We seek the value $x_0$ such that the regression curve attains its maximum at $X = x_0$. This point is given by $\hat{x}_0 = 2.0751038$.

**d** Let $x_c$ represent the elapsed time at which the drug concentration in the bloodstream falls below 1 microgram/deciliter (after attaining the peak concentration). We obtain $\hat{x}_c = 4.402935787$.

**9.4.1** **a** The linearizing transformation is $Z = 1/Y$.

**b** $\hat{\beta}_1 = 1.4026$, $\hat{\beta}_2 = -1.2903$, $\hat{\beta}_3 = 0.27640$

**9.4.2** **a** The linearizing transformation is $Z = -\ln\left[-\ln\left(1 - \frac{Y}{2}\right)\right]$.

**b** The estimates of $\beta_2$ and $\beta_3$ obtained after linearization are $\hat{\beta}_2 = -1.081$ and $\hat{\beta}_3 = 13.0413$.

9.5.1 **b** $\hat{\beta}_1 = 0.00656892$, $\hat{\beta}_2 = -0.00028246$, $\hat{\beta}_3 = 0.00000749$, $\hat{\beta}_4 = -80.65213231$

9.5.2 **b** $\hat{\beta}_1 = 0.995$, $\hat{\beta}_2 = -115.929282$, $\hat{\beta}_3 = 5402.347473$, $\hat{\beta}_4 = 0.009437$
Also $\hat{\sigma} = 0.0145530$.

**d** The required value of $x_0$ is obtained to be .01949075756.

**e** The required quantity is equal to $\theta$ where $\theta = \frac{\beta_1}{[1 + e^{-\beta_2}]^{\beta_4}}$.
$C[0.1786 \le \theta \le 0.5464] \ge 0.85$ (approximately).