## Chapter 11 Selected Problem Solutions

Section 11-2

11-1. a) $y_i = \beta_0 + \beta_1 x_1 + \varepsilon_i$

$$S_{xx} = 157.42 - \frac{43^2}{14}$$
$$= 25.348571$$
$$S_{xy} = 1697.80 - \frac{43(572)}{14}$$
$$= -59.057143$$
$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{-59.057143}{25.348571} = -2.330$$
$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x} = \frac{572}{14} - (-2.3298017)(\frac{43}{14}) = 48.013$$

b) $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

$$\hat{y} = 48.012962 - 2.3298017(4.3) = 37.99$$

c) $\hat{y} = 48.012962 - 2.3298017(3.7) = 39.39$

d) $e = y - \hat{y} = 46.1 - 39.39 = 6.71$

11-5. a)

```
Regression Analysis - Linear model: Y = a+bX
Dependent variable: SalePrice                Independent variable: Taxes
--------------------------------------------------------------------------------
                                Standard             T              Prob.
Parameter       Estimate         Error             Value            Level
Intercept        13.3202         2.57172          5.17948           .00003
Slope            3.32437        0.390276            8.518           .00000
--------------------------------------------------------------------------------
                            Analysis of Variance
Source              Sum of Squares     Df  Mean Square    F-Ratio  Prob. Level
Model                    636.15569      1    636.15569    72.5563       .00000
Residual                 192.89056     22      8.76775
--------------------------------------------------------------------------------
Total (Corr.)            829.04625     23
Correlation Coefficient = 0.875976          R-squared =  76.73 percent
Stnd. Error of Est. = 2.96104
```

$$\hat{\sigma}^2 = 8.76775$$

If the calculations were to be done by hand use Equations (11-7) and (11-8).

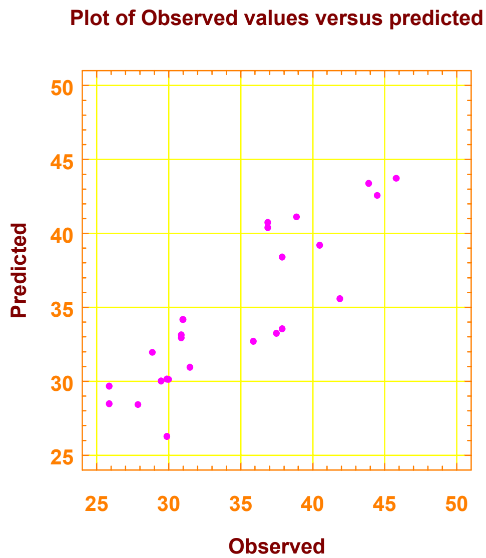
$\hat{y} = 13.3202 + 3.32437x$

b) $\hat{y} = 13.3202 + 3.32437(7.5) = 38.253$

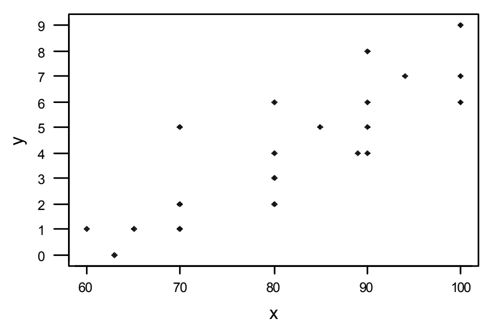
c) $\hat{y} = 13.3202 + 3.32437(5.8980) = 32.9273$

$$\hat{y} = 32.9273$$
$$e = y - \hat{y} = 30.9 - 32.9273 = -2.0273$$

d) All the points would lie along the 45% axis line. That is, the regression model would estimate the values exactly. At this point, the graph of observed vs. predicted indicates that the simple linear regression model provides a reasonable fit to the data.




11-9. a) Yes, a linear regression would seem appropriate, but one or two points appear to be outliers.

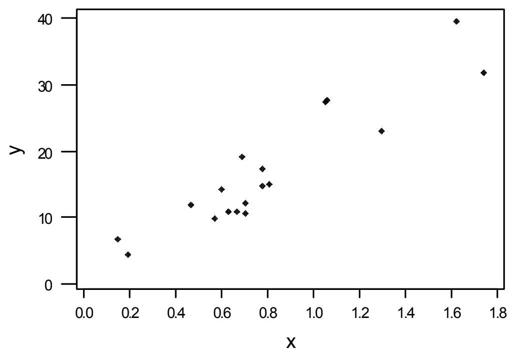



```
Predictor          Coef      SE Coef          T        P
Constant         -9.813        2.135      -4.60    0.000
x               0.17148      0.02566       6.68    0.000

S = 1.408        R-Sq = 71.3%      R-Sq(adj) = 69.7%

Analysis of Variance

Source            DF          SS          MS          F         P
Regression         1      88.520      88.520      44.66     0.000
Residual Error    18      35.680       1.982
Total             19     124.200
```

b) $\hat{\sigma}^2 = 1.9818$ and $\hat{y} = -9.8131 + 0.171484x$

c) $\hat{y} = 4.76301$ at x = 85

11-11. a) Yes, a linear regression would seem appropriate.




| Predictor | Coef | SE Coef | T | P |
|---|---|---|---|---|
| Constant | 0.470 | 1.936 | 0.24 | 0.811 |
| x | 20.567 | 2.142 | 9.60 | 0.000 |

S = 3.716     R-Sq = 85.2%     R-Sq(adj) = 84.3%

Analysis of Variance

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 1 | 1273.5 | 1273.5 | 92.22 | 0.000 |
| Residual Error | 16 | 220.9 | 13.8 | | |
| Total | 17 | 1494.5 | | | |

b) $\hat{\sigma}^2 = 13.81$

$\hat{y} = 0.470467 + 20.5673x$

c) $\hat{y} = 0.470467 + 20.5673(1) = 21.038$

d) $\hat{y} = 10.1371 \quad e = 1.6629$

## Section 11-4

11-21. Refer to ANOVA of Exercise 11-5

a) 1) The parameter of interest is the regressor variable coefficient, $\beta_1$.

2) $H_0 : \beta_1 = 0$

3) $H_1 : \beta_1 \neq 0$

4) $\alpha = 0.05$, using t-test

5) The test statistic is $t_0 = \dfrac{\hat{\beta}_1}{se(\hat{\beta}_1)}$

6) Reject $H_0$ if $t_0 < -t_{\alpha/2,n-2}$ where $-t_{0.025,22} = -2.074$ or $t_0 > t_{0.025,22} = 2.074$

7) Using the results from Exercise 11-5

$$t_0 = \frac{3.32437}{0.390276} = 8.518$$

8) Since 8.518 > 2.074 reject $H_0$ and conclude the model is useful $\alpha = 0.05$.

b) 1) The parameter of interest is the slope, $\beta_1$

2) $H_0:\beta_1 = 0$

3) $H_1:\beta_1 \neq 0$

4) $\alpha = 0.05$

5) The test statistic is $f_0 = \frac{MS_R}{MS_E} = \frac{SS_R / 1}{SS_E / (n-2)}$

6) Reject $H_0$ if $f_0 > f_{\alpha,1,22}$ where $f_{0.01,1,22} = 4.303$

7) Using the results from Exercise 10-5

$$f_0 = \frac{636.15569 / 1}{192.89056 / 22} = 72.5563$$

8) Since 72.5563 > 4.303, reject $H_0$ and conclude the model is useful at a significance $\alpha = 0.05$.

The F-statistic is the square of the t-statistic. The F-test is a restricted to a two-sided test, whereas the t-test could be used for one-sided alternative hypotheses.

c) $se(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} = \sqrt{\frac{8.7675}{57.5631}} = .39027$

$$se(\hat{\beta}_0) = \sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{\bar{x}}{S_{xx}}\right]} = \sqrt{8.7675\left[\frac{1}{24} + \frac{6.4049^2}{57.5631}\right]} = 2.5717$$

d) 1) The parameter of interest is the intercept, $\beta_0$.

2) $H_0 : \beta_0 = 0$

3) $H_1 : \beta_0 \neq 0$

4) $\alpha = 0.05$, using t-test

5) The test statistic is $t_0 = \frac{\hat{\beta}_0}{se(\hat{\beta}_0)}$

6) Reject $H_0$ if $t_0 < -t_{\alpha/2,n-2}$ where $-t_{0.025,22} = -2.074$ or $t_0 > t_{0.025,22} = 2.074$

7) Using the results from Exercise 11-5

$$t_0 = \frac{13.3201}{2.5717} = 5.2774$$

8) Since 5.2774 > 2.074 reject $H_0$ and conclude the intercept is not zero at $\alpha = 0.05$.

11-25. Refer to ANOVA of Exercise 11-9

a) $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

$\alpha = 0.05$

$f_0 = 44.6567$

$f_{.05,1,18} = 4.416$

$f_0 > f_{\alpha,1,18}$

Therefore, reject $H_0$. P-value = 0.000003.

b) $se(\hat{\beta}_1) = 0.0256613$

$se(\hat{\beta}_0) = 2.13526$

c) $H_0 : \beta_0 = 0$

$H_1 : \beta_0 \neq 0$

$\alpha = 0.05$

$$t_0 = -4.59573$$
$$t_{.025,18} = 2.101$$
$$|t_0| > t_{\alpha/2,18}$$

Therefore, reject $H_0$. P-value = 0.00022.

## Sections 11-5 and 11-6

11-31. $t_{\alpha/2,n-2} = t_{0.025,12} = 2.179$

a) 95% confidence interval on $\beta_1$.

$$\hat{\beta}_1 \pm t_{\alpha/2,n-2}se(\hat{\beta}_1)$$
$$-2.3298 \pm t_{.025,12}(0.2697)$$
$$-2.3298 \pm 2.179(0.2697)$$
$$-2.9175. \le \beta_1 \le -1.7421.$$

b) 95% confidence interval on $\beta_0$.

$$\hat{\beta}_0 \pm t_{.025,12}se(\hat{\beta}_0)$$
$$48.0130 \pm 2.179(0.5959)$$
$$46.7145 \le \beta_0 \le 49.3114.$$

c) 95% confidence interval on $\mu$ when $x_0 = 2.5$.

$$\hat{\mu}_{Y|x_0} = 48.0130 - 2.3298(2.5) = 42.1885$$
$$\hat{\mu}_{Y|x_0} \pm t_{.025,12}\sqrt{\hat{\sigma}^2(\tfrac{1}{n} + \tfrac{(x_0-\bar{x})^2}{S_{xx}})}$$
$$42.1885 \pm (2.179)\sqrt{1.844(\tfrac{1}{14} + \tfrac{(2.5-3.0714)^2}{25.3486})}$$
$$42.1885 \pm 2.179(0.3943)$$
$$41.3293 \le \hat{\mu}_{Y|x_0} \le 43.0477$$

d) 95% on prediction interval when $x_0 = 2.5$.

$$\hat{y}_0 \pm t_{.025,12}\sqrt{\hat{\sigma}^2(1 + \tfrac{1}{n} + \tfrac{(x_0-\bar{x})^2}{S_{xx}})}$$
$$42.1885 \pm 2.179\sqrt{1.844(1 + \tfrac{1}{14} + \tfrac{(2.5-3.0714)^2}{25.348571})}$$
$$42.1885 \pm 2.179(1.1808)$$
$$38.2489 \le y_0 \le 46.1281$$

It is wider because it depends on both the error associated with the fitted model as well as that with the future observation.

11-35. 99 percent confidence intervals for coefficient estimates

| | Estimate | Standard error | Lower Limit | Upper Limit |
|---|---|---|---|---|
| CONSTANT | -6.33550 | 1.66765 | -11.6219 | -1.05011 |
| Temperature | 9.20836 | 0.03377 | 9.10130 | 9.93154 |

a) $9.10130 \le \beta_1 \le 9.31543$

b) $-11.6219 \le \beta_0 \le -1.04911$

c) $500.124 \pm (2.228)\sqrt{3.774609(\tfrac{1}{12} + \tfrac{(55-46.5)^2}{3308.9994})}$

$$500.124 \pm 1.4037586$$
$$498.72024 \le \hat{\mu}_{Y|x_0} \le 501.52776$$

d) $500.124 \pm (2.228)\sqrt{3.774609(1 + \tfrac{1}{12} + \tfrac{(55-46.5)^2}{3308.9994})}$

$$500.124 \pm 4.5505644$$

$$495.57344 \le y_0 \le 504.67456$$

It is wider because the prediction interval includes error for both the fitted model and from that associated with the future observation.

11-41 a) $-43.1964 \le \beta_1 \le -30.7272$

b) $2530.09 \le \beta_0 \le 2720.68$

c) $1886.154 \pm (2.101)\sqrt{9811.21\left(\frac{1}{20} + \frac{(20-13.3375)^2}{1114.6618}\right)}$

$1886.154 \pm 62.370688$

$1823.7833 \le \mu_{y|x_0} \le 1948.5247$

d) $1886.154 \pm (2.101)\sqrt{9811.21\left(1 + \frac{1}{20} + \frac{(20-13.3375)^2}{1114.6618}\right)}$

$1886.154 \pm 217.25275$

$1668.9013 \le y_0 \le 2103.4067$

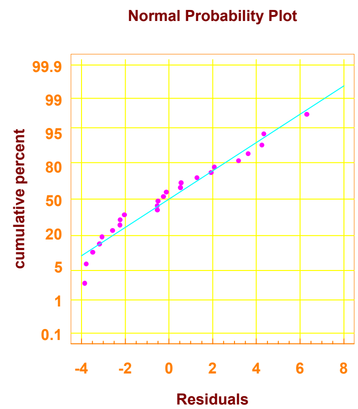
## Section 11-7

11-43. Use the Results of exercise 11-5 to answer the following questions.

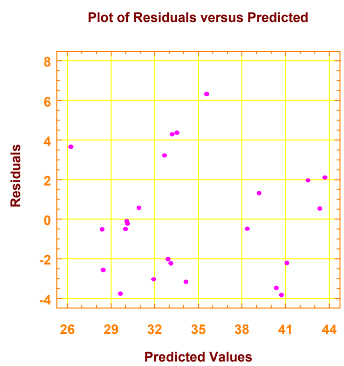
a)

| SalePrice | Taxes | Predicted | Residuals |
|---|---|---|---|
| 25.9 | 4.9176 | 29.6681073 | -3.76810726 |
| 29.5 | 5.0208 | 30.0111824 | -0.51118237 |
| 27.9 | 4.5429 | 28.4224654 | -0.52246536 |
| 25.9 | 4.5573 | 28.4703363 | -2.57033630 |
| 29.9 | 5.0597 | 30.1405004 | -0.24050041 |
| 29.9 | 3.8910 | 26.2553078 | 3.64469225 |
| 30.9 | 5.8980 | 32.9273208 | -2.02732082 |
| 28.9 | 5.6039 | 31.9496232 | -3.04962324 |
| 35.9 | 5.8282 | 32.6952797 | 3.20472030 |
| 31.5 | 5.3003 | 30.9403441 | 0.55965587 |
| 31.0 | 6.2712 | 34.1679762 | -3.16797616 |
| 30.9 | 5.9592 | 33.1307723 | -2.23077234 |
| 30.0 | 5.0500 | 30.1082540 | -0.10825401 |
| 36.9 | 8.2464 | 40.7342742 | -3.83427422 |
| 41.9 | 6.6969 | 35.5831610 | 6.31683901 |
| 40.5 | 7.7841 | 39.1974174 | 1.30258260 |
| 43.9 | 9.0384 | 43.3671762 | 0.53282376 |
| 37.5 | 5.9894 | 33.2311683 | 4.26883165 |
| 37.9 | 7.5422 | 38.3932520 | -0.49325200 |
| 44.5 | 8.7951 | 42.5583567 | 1.94164328 |
| 37.9 | 6.0831 | 33.5426619 | 4.35733807 |
| 38.9 | 8.3607 | 41.1142499 | -2.21424985 |
| 36.9 | 8.1400 | 40.3805611 | -3.48056112 |
| 45.8 | 9.1416 | 43.7102513 | 2.08974865 |

b) Assumption of normality does not seem to be violated since the data appear to fall along a straight line.




c) No serious departure from assumption of constant variance. This is evident by the random pattern of the residuals.

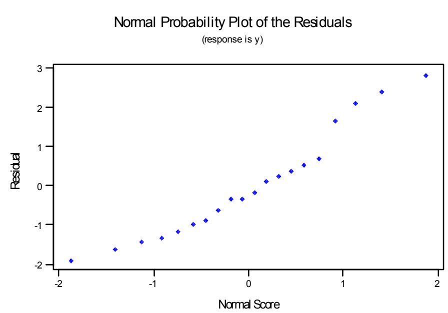




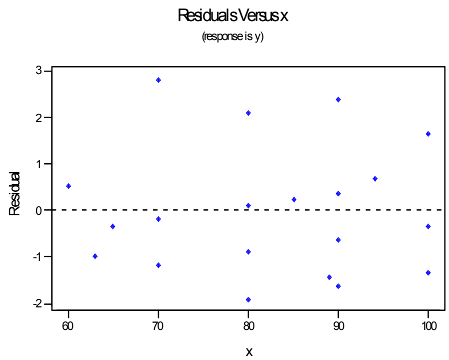
d) $R^2 \equiv 76.73\%$ ;

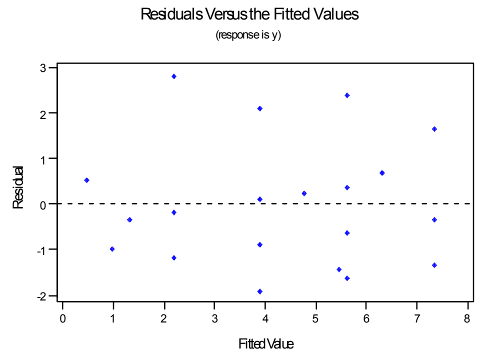
11-47. a) $R^2 = 71.27\%$

b) No major departure from normality assumptions.



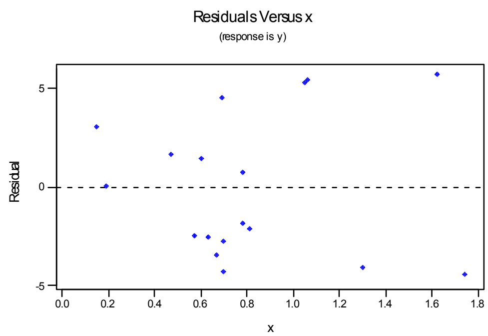

c) Assumption of constant variance appears reasonable.

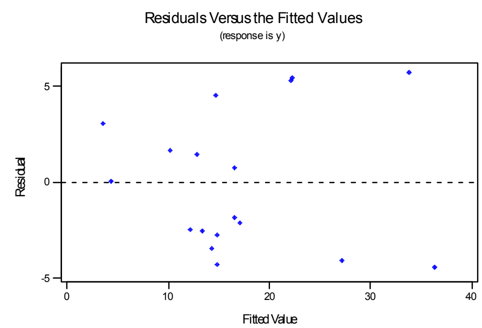


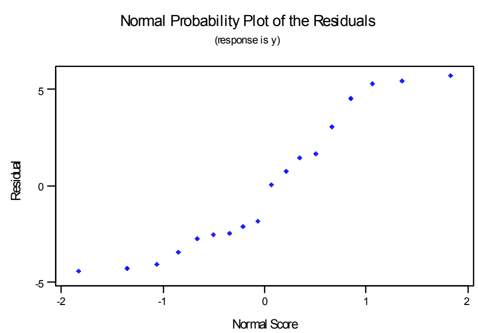

11-49. a) $R^2 = 85.22\%$

b) Assumptions appear reasonable, but there is a suggestion that variability increases with $\hat{y}$.




c) Normality assumption may be questionable. There is some " bending" away from a straight line in the tails of the normal probability plot.




## Section 11-10

11-55. a) $\hat{y} = -0.0280411 + 0.990987x$

b) $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 79.838$

$f_{.05,1,18} = 4.41$

$f_0 >> f_{\alpha,1,18}$

Reject $H_0$.

c) $r = \sqrt{0.816} = 0.903$

d) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$ $\alpha = 0.05$

$$t_0 = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}} = \frac{0.90334\ \sqrt{18}}{\sqrt{1-0.816}} = 8.9345$$

$t_{.025,18} = 2.101$

$t_0 > t_{\alpha/2,18}$

Reject $H_0$.

e) $H_0 : \rho = 0.5$

$H_1 : \rho \neq 0.5$ $\alpha = 0.05$

$z_0 = 3.879$

$z_{.025} = 1.96$

$z_0 > z_{\alpha/2}$

Reject $H_0$.

f) $\tanh(\text{arctanh } 0.90334 - \frac{z_{.025}}{\sqrt{17}}) \leq \rho \leq \tanh(\text{arctanh } 0.90334 + \frac{z_{.025}}{\sqrt{17}})$ where $z_{.025} = 1.96$.

$0.7677 \leq \rho \leq 0.9615$.

11-59 n = 50 r = 0.62

a) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$ $\alpha = 0.01$

$$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.62\sqrt{48}}{\sqrt{1-(0.62)^2}} = 5.475$$

$t_{.005,48} = 2.683$

$t_0 > t_{0.005,48}$

Reject $H_0$. P-value $\cong 0$

b) $\tanh(\text{arctanh } 0.62 - \frac{z_{.005}}{\sqrt{47}}) \leq \rho \leq \tanh(\text{arctanh } 0.62 + \frac{z_{.005}}{\sqrt{47}})$ where $z_{.005} = 2.575$.

$0.3358 \leq \rho \leq 0.8007$.

c) Yes.

11-61. a) $r = 0.933203$

a) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$ $\alpha = 0.05$

$$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.933203\ \sqrt{15}}{\sqrt{1-(0.8709)}} = 10.06$$

$t_{.025,15} = 2.131$

$t_0 > t_{\alpha/2,15}$

Reject $H_0$.

c) $\hat{y} = 0.72538 + 0.498081x$

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 101.16$

$f_{.05,1,15} = 4.545$

$f_0 >> f_{\alpha,1,15}$

Reject $H_0$. Conclude that the model is significant at $\alpha = 0.05$. This test and the one in part b are identical.

d) $H_0 : \beta_0 = 0$

$H_1 : \beta_0 \neq 0$ $\alpha = 0.05$

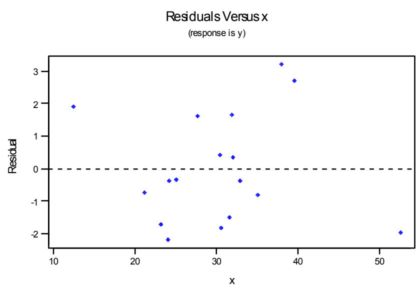
$t_0 = 0.468345$

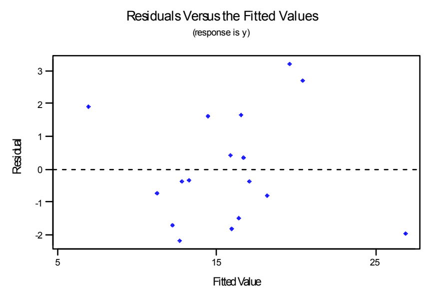
$t_{.025,15} = 2.131$

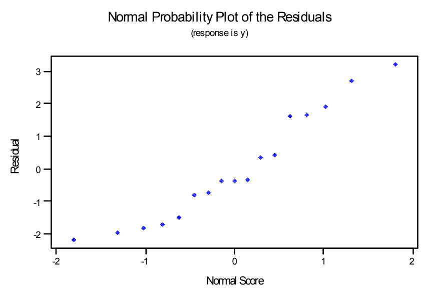
$t_0 \not> t_{\alpha/2,15}$

Do not reject $H_0$. We cannot conclude $\beta_0$ is different from zero.

e) No serious problems with model assumptions are noted.




Supplemental

11-65. a) $\hat{y} = 93.34 + 15.64x$

b) $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 12.872$

$f_{0.05,1,14} = 4.60$

$f_0 > f_{0.05,1,14}$

Reject $H_0$. Conclude that $\beta_1 \neq 0$ at $\alpha = 0.05$.

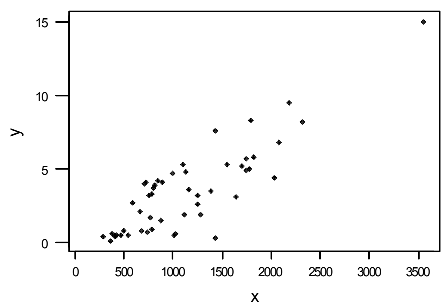
c) $(7.961 \leq \beta_1 \leq 23.322)$

d) $(74.758 \leq \beta_0 \leq 111.923)$

e) $\hat{y} = 93.34 + 15.64(2.5) = 132.44$

$$132.44 \pm 2.145\sqrt{136.27\left[\frac{1}{16} + \frac{(2.5-2.325)^2}{7.017}\right]}$$
$$132.44 \pm 6.26$$
$$126.18 \le \hat{\mu}_{Y|x_0=2.5} \le 138.70$$

11-67 a)




b) $\hat{y} = -0.8819 + 0.00385x$

c) $H_0 : \beta_1 = 0$

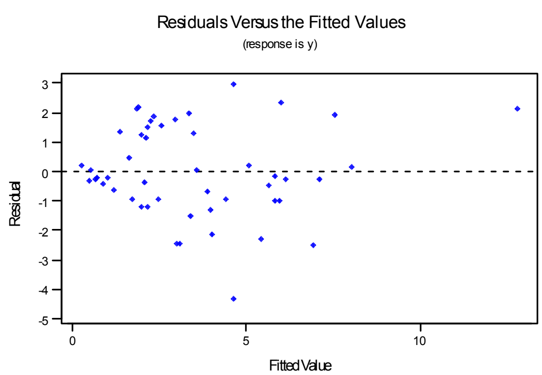
$H_1 : \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 122.03$

$f_0 > f_{\alpha,1,48}$

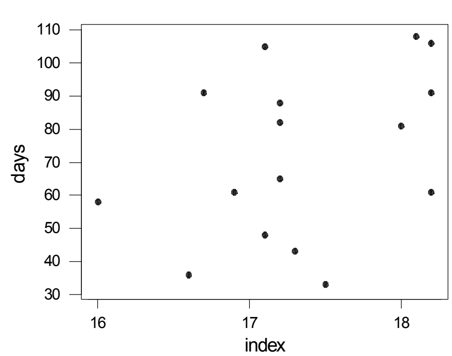
Reject $H_0$ . Conclude that regression model is significant at $\alpha = 0.05$

d) No, it seems the variance is not constant, there is a funnel shape.




e) $\hat{y}^* = 0.5967 + 0.00097x$ . Yes, the transformation stabilizes the variance.

11-71 a)



b) The regression equation is

$\hat{y} = -193 + 15.296x$

| Predictor | Coef | SE Coef | T | P |
|---|---|---|---|---|
| Constant | -193.0 | 163.5 | -1.18 | 0.258 |
| x | 15.296 | 9.421 | 1.62 | 0.127 |

S = 23.79     R-Sq = 15.8%     R-Sq(adj) = 9.8%

Analysis of Variance

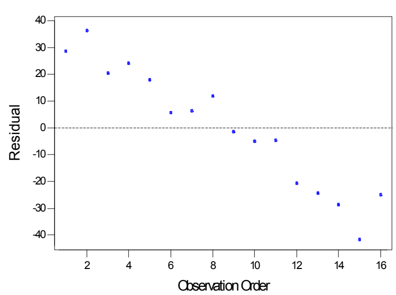
| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 1 | 1492.6 | 1492.6 | 2.64 | 0.127 |
| Error | 14 | 7926.8 | 566.2 | | |
| Total | 15 | 9419.4 | | | |

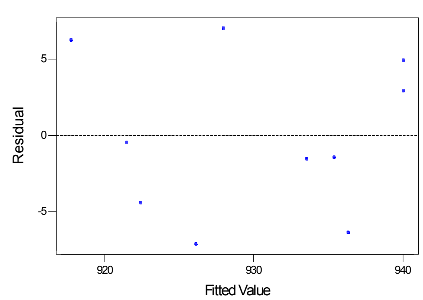
Cannot reject $H_o$; therefore we conclude that the model is not significant. Therefore the seasonal meteorological index (x) is not a reliable predictor of the number of days that the ozone level exceeds 0.20 *ppm* (*y*).

c) 95% CI on $\beta_1$

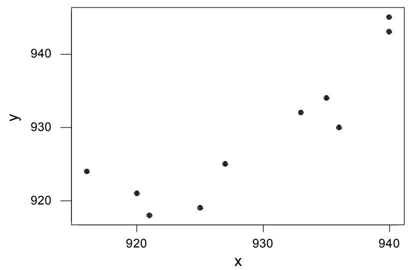
$$\hat{\beta}_1 \pm t_{\alpha/2,n-2}se(\hat{\beta}_1)$$
$$15.296 \pm t_{.025,12}(9.421)$$
$$15.296 \pm 2.145(9.421)$$
$$-4.912 \le \beta_1 \le 35.504$$

d) The normality plot of the residuals is satisfactory. However, the plot of residuals versus run order exhibits a strong downward trend. This could indicate that there is another variable should be included in the model, one that changes with time.



11-75 a)




b) $\hat{y} = 33.3 + 0.9636x$

c)

```
Predictor          Coef      SE Coef          T        P
Constant           33.3        171.7       0.19    0.851
x                0.9639       0.1848       5.22    0.001

S = 4.805       R-Sq = 77.3%     R-Sq(adj) = 74.4%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         1      628.18      628.18     27.21    0.001
Residual Error     8      184.72       23.09
Total              9      812.90
```

Reject the hull hypothesis and conclude that the model is significant. 77.3% of the variability is explained by the model.

d) $H_0 : \beta_1 = 1$

$H_1 : \beta_1 \neq 1$ α=.05

$$t_0 = \frac{\hat{\beta}_1 - 1}{se(\hat{\beta}_1)} = \frac{0.9639 - 1}{0.1848} = -0.1953$$

$$t_{a/2,n-2} = t_{.025,8} = 2.306$$

Since $t_0 > -t_{a/2,n-2}$, we cannot reject $H_o$ and we conclude that there is not enough evidence to reject the claim that the devices produce different temperature measurements. Therefore, we assume the devices produce equivalent measurements.

e) The residual plots to not reveal any major problems.