CHAPTER 4

4.1

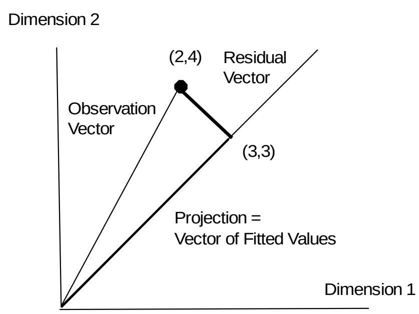
$$X'X = \begin{bmatrix} 10 & 55 \\ 55 & 385 \end{bmatrix};\ (X'X)^{-1} = \begin{bmatrix} 0.4667 & -0.0667 \\ -0.0667 & 0.0121 \end{bmatrix};$$

$$V(\hat{\boldsymbol{\beta}}) = \sigma^2 \begin{bmatrix} 0.4667 & -0.0667 \\ -0.0667 & 0.0121 \end{bmatrix}$$

$$V(\hat{\beta}_0) = (0.4667)\sigma^2;\ V(\hat{\beta}_1) = (0.0121)\sigma^2$$

4.2 L(**1**) represents the 45 degree line through the origin in two-dimensional space. Projecting the observation vector $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ onto the subspace L(**1**) results in the fitted values $\hat{\boldsymbol{\mu}} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and the least squares estimate $\hat{\beta}_0 = 3$. The residual vector $\boldsymbol{e} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} - \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and the projection $\hat{\boldsymbol{\mu}} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ are orthogonal. The picture is given below.

Exercise 4.2



4.3 L(X) is the two-dimensional subspace in three-dimensional space that is described by all linear combinations of the two vectors, $\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\boldsymbol{x} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$. You need to visualize this as a plane in three-dimensional space. The orthogonal projection of the observation vector $\boldsymbol{y} = \begin{bmatrix} 2.2 \\ 3.9 \\ 3.1 \end{bmatrix}$ onto this plane results in the vector of fitted values (the projection) $\hat{\boldsymbol{\mu}} = \begin{bmatrix} 2.21667 \\ 3.91667 \\ 3.06667 \end{bmatrix}$ and the least squares estimates $\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} 1.36667 \\ 0.85000 \end{bmatrix}$, satisfying $\hat{\boldsymbol{\mu}} = \begin{bmatrix} 2.21667 \\ 3.91667 \\ 3.06667 \end{bmatrix} = \mathrm{X}\hat{\boldsymbol{\beta}} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1.36667 \\ 0.85000 \end{bmatrix}$. The residual vector $\boldsymbol{e} = \begin{bmatrix} 2.1 - 2.21667 \\ 3.9 - 3.91667 \\ 3.1 - 3.06667 \end{bmatrix} = \begin{bmatrix} -0.01667 \\ -0.01667 \\ 0.03333 \end{bmatrix}$ and the projection $\hat{\boldsymbol{\mu}} = \begin{bmatrix} 2.21667 \\ 3.91667 \\ 3.06667 \end{bmatrix}$ are orthogonal.

The difference of the data vector $\boldsymbol{y}$ and the projection $\hat{\boldsymbol{\mu}}$ is quite small, indicating that the data vector is almost in the plane spanned by the matrix X.

4.4 L(X) is the two-dimensional subspace in three-dimensional space that is described by all linear combinations of the two vectors, $\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\boldsymbol{x} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$. You need to visualize this as a plane in three-dimensional space. The orthogonal projection of the observation vector $\boldsymbol{y} = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$ onto this plane results in the vector of fitted values (the projection) $\hat{\boldsymbol{\mu}} = \begin{bmatrix} 3 \\ 5 \\ 4 \end{bmatrix}$ and the least squares estimates $\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, satisfying

$\hat{\boldsymbol{\mu}} = \begin{bmatrix} 3 \\ 5 \\ 4 \end{bmatrix} = X\hat{\boldsymbol{\beta}} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$. The residual vector $\boldsymbol{e} = \begin{bmatrix} 2-3 \\ 4-5 \\ 6-4 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$ and the projection $\hat{\boldsymbol{\mu}} = \begin{bmatrix} 3 \\ 5 \\ 4 \end{bmatrix}$ are orthogonal.

Note that the difference of the data vector and the projection is larger here than in Exercise 4.3. The data vector is not close to the space spanned by the matrix X.

4.5

(a) $V(\hat{\beta}_1) = 18$

(b) $\text{Cov}(\hat{\beta}_1, \hat{\beta}_3) = 1.2$

(c) $\text{Corr}(\hat{\beta}_1, \hat{\beta}_3) = 0.0943$

(d) $V(\hat{\beta}_1 - \hat{\beta}_3) = V(\hat{\beta}_1) + V(\hat{\beta}_3) - 2\text{Cov}(\hat{\beta}_1, \hat{\beta}_3) = 24.6$

4.6

(a) $V(\hat{\beta}_2) = 4$; $\text{s.e.}(\hat{\beta}_2) = 2$

(b) $t(\hat{\beta}_2) = \hat{\beta}_2 / \text{s.e.}(\hat{\beta}_2) = 15/2 = 7.5$; p-value = $2P(t(12) \geq 7.5) < 0.001$; reject $\beta_2 = 0$ in favor of $\beta_2 \neq 0$

(c) $\text{Cov}(\hat{\beta}_1, \hat{\beta}_2) = -0.05$

(d) Test $\beta_1 - \beta_2 = 0$; $V(\hat{\beta}_1 - \hat{\beta}_2) = V(\hat{\beta}_1) + V(\hat{\beta}_2) - 2\text{Cov}(\hat{\beta}_1, \hat{\beta}_2) = 1 + 4 + 1 = 6$;
$\hat{\beta}_1 - \hat{\beta}_2 / \text{s.e.}(\hat{\beta}_1 - \hat{\beta}_2) = -3/\sqrt{6} = -1.22$; p-value = $2P(t(12) \leq -1.22) = 0.136$;
conclude $\beta_1 - \beta_2 = 0$, or $\beta_1 = \beta_2$.
95% confidence interval for $\beta_1 - \beta_2$: $-3 \pm (2.179)\sqrt{6}$ or (-8.34, 2.34) covers zero.

(e) SST = 120, SSE = 2(15-3) = 24, and SSR = 96; F = (96/2)/(24/12) = 24; very small p-value; reject $\beta_1 - \beta_2 = 0$.

4.7

(a) $R^2 = 0.9324$

(b) F-statistic = 110.35; p-value = 0.000; reject $\beta_1 = \beta_2 = \beta_3 = 0$

(c) 95% confidence interval for β_{taxes}: (0.074, 0.306); reject $\beta_{taxes} = 0$; cannot simplify model
95% confidence Interval for β_{baths}: (-16.83, 180.57); can not reject $\beta_{baths} = 0$; can simplify model by dropping "baths"

4.8

(a) R^2 = 500074/541119 = 0.9241

(b) F-statistic = 152.29; p-value = 0.000; reject $\beta_1 = \beta_2 = 0$

(c) t-ratio for taxes $= 0.24237/0.04884 = 4.96$; p-value = $P(t(25) \geq 4.96) = 0.0000$

reject $\beta_{\text{taxes}} = 0$; response is related to taxes.

4.9

(a) From $\hat{\beta} = (X'X)^{-1}X'y$, we obtain $\hat{\beta}_0 = 885.161, \hat{\beta}_1 = -6.571, \hat{\beta}_2 = -1.374$;

s = 36.49;

From $V(\hat{\beta}) = s^2(X'X)^{-1}$: $\text{s.e.}(\hat{\beta}_0) = 61.75, \text{s.e.}(\hat{\beta}_1) = 0.5832, \text{s.e.}(\hat{\beta}_2) = 0.1943$

(b) $t(\hat{\beta}_0) = 14.33, t(\hat{\beta}_1) = -11.27, t(\hat{\beta}_2) = -7.07$; 97.5th percentile: t(27,0.975) = 2.052;

can reject $\beta_1 = 0$; can reject $\beta_2 = 0$.

(c) R^2 = 0.841

4.10

(a) Estimated equation: $\hat{\mu} = 3.453 + 0.496x_1 + 0.0092x_2$; s^2 = 4.7403;

s.e.($\hat{\beta}_0$) = 2.431, s.e.($\hat{\beta}_1$) = 0.00605, s.e.($\hat{\beta}_2$) = 0.00097

(b) $t(\hat{\beta}_1) = 0.496/0.00605 = 81.89$; p-value (2-sided) = $2\,P(t(12) > 81.89)$ = 0.000, which is very small. We reject the null hypothesis $\beta_1 = 0$.

$t(\hat{\beta}_2) = 0.009191/0.00097 = 9.49$; p-value (2-sided) = 0.000, which is very small. We reject the null hypothesis $\beta_2 = 0$.

Neither of the two explanatory variables can be omitted from the model.

4.11

(a) Minitab output:

```
The regression equation is
Y = 295 - 481 X1 - 829 X2 + 0.00794 X3 + 2.36 X4

Predictor        Coef     SE Coef          T        P
Constant       295.33       40.18       7.35    0.000
X1             -480.8       150.4      -3.20    0.006
X2             -829.4       196.5      -4.22    0.001
X3           0.007936    0.003554       2.23    0.041
X4             2.3603      0.7616       3.10    0.007

S = 46.77       R-Sq = 88.3%     R-Sq(adj) = 85.1%
```

```
Analysis of Variance

Source           DF        SS         MS       F        P
Regression        4    246538      61635   28.18    0.000
Residual Error   15     32807       2187
Total            19    279345
```

(b) Test $\beta_1 = 0 : t(\hat{\beta}_1) = -3.20$; p-value = 0.006; reject $\beta_1 = 0$; the number of beds in for profit hospitals is important.

(c) The observations (for the two time periods for each state) look very similar and, most likely, they are correlated. If the correlation is very high, it is reasonable to discard one of them or average the two observations, and reanalyze the data.

(d) Instead of selecting ten states at random, one could classify the states into three groups according to population size - small, medium, and large - and select three or four hospitals at random from each category.

4.12 The output from R software, using the function lm(formula = usage ~ TEMP + PROD + DAYS + PAYR + HOUR) is given below:

```
Coefficients:
              Estimate    Std. Error    t value      Pr (>|t|)
(Intercept)  39.437054    12.110986      3.256       0.00765
TEMP          0.084067     0.060469      1.390       0.19194
PROD          0.001876     0.000607      3.091       0.01027
DAYS          0.131704     0.289800      0.454       0.65833
PAYS         -0.215677     0.098810     -2.183       0.05162
HOUR         -0.014475     0.030052     -0.482       0.63949

Residual standard error: 3.213 on 11 degrees of freedom
Multiple R-Squared: 0.6446, Adjusted R-squared: 0.4831
F-statistic: 3.991 on 5 and 11 DF, p-value: 0.02607
```

R^2 = 0.6446, and the regression model is significant at 2.6% level. The output indicates that PROD is significant at the 1% level, even if other variables are present in the model. PAYS is also marginally significant (p-value = 0.051). All other variables are not significant when added last to the model. The model can be simplified

(b) In order to test $\beta_1 = \beta_3 = \beta_5 = 0$, we need to fit a reduced model that includes just x_2 and x_4. The R output for the reduced model with lm(formula = USAGE ~ PROD + PAYR) is listed below

```
Coefficients:
                 Estimate  Std. Error t value  Pr (>|t|)
(Intercept)    46.0177241  10.1085905   4.552  0.000452
PROD            0.0020353   0.0005587   3.643  0.002663
PAYR           -0.2157919   0.0895867  -2.409  0.030356

Residual standard error: 3.117 on 14 degrees of freedom
Multiple R-Squared: 0.5743, Adjusted R-squared: 0.5135
F-statistic: 9.442 on 2 and 14 DF, p-value: 0.002535
```

The additional sum of squares = ResidualSS (reduced model) – ResidualSS (full model) = SSR(full model) – SSR(reduced model) = 205.956 –183.48 and
F = [(205.956 –183.48)/2]/(3.213)^2 = 1.09; p-value = P(F(2,11) > 1.09) = 0.37; we can not reject $\beta_1 = \beta_3 = \beta_5 = 0$.

(c) We prefer the reduced model $\hat{\mu} = 46.02 + 0.00204\text{PROD} - 0.216\text{PAYR}$; R^2 = 0.574 (only slightly smaller than the R^2 of the full model = 0.6446).

(d) Production has the smallest p-value.

(e) Water usage as linear function of PROD and PAYR. For fixed value of PAYR, each unit increase in production increases water use by 0.0020353 (gallons/100). Similarly, for a fixed value of PROD, a unit increase in PAYR decreases water usage by 0.2157919 (gallons/100).

4.13

(a) Minitab output:

```
The regression equation is
Y = 177 + 2.17 X1 + 3.54 X2 - 22.2 X3 + 0.204 X4

Predictor        Coef     SE Coef          T        P
Constant      177.229       8.787      20.17    0.000
X1             2.1702      0.6737       3.22    0.009
X2             3.5380      0.1092      32.41    0.000
X3           -22.1583      0.5454     -40.63    0.000
X4             0.2035      0.3189       0.64    0.538

S = 5.119       R-Sq = 99.7%     R-Sq(adj) = 99.6%

Analysis of Variance

Source            DF          SS          MS         F         P
Regression         4       89285       22321    851.72     0.000
Residual Error    10         262          26
Total             14       89547
```

(b) R^2 = 0.997; estimates are part of the output given above

(c)

(i) t-ratio = 0.64; p-value = 0.538; conclude $\beta_4 = 0$

(ii) F = [(43968 – 262)/2]/[262/10] = 834.1; p-value = P(F(2,10) > 834.1) = 0.0000; reject $\beta_3 = \beta_4 = 0$

(iii) F = (58575 – 262)/(262/10) = 2,225.7; p-value = P(F(2,10) > 2,225.7) = 0.0000; reject $\beta_2 = \beta_3$

```
The regression equation is
Y = - 61 + 4.61 X1 + 3.05 X2+X3 + 2.54 X4

Predictor        Coef     SE Coef          T        P
Constant        -61.4       102.4      -0.60    0.561
X1              4.613       9.575       0.48    0.639
X2+X3           3.051       1.549       1.97    0.075
X4              2.541       4.490       0.57    0.583

S = 72.97       R-Sq = 34.6%     R-Sq(adj) = 16.7%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         3       30972       10324      1.94    0.182
Residual Error    11       58575        5325
Total             14       89547
```

(iv) F = 851.72; p-value = 0.0000; reject $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

(d) Minitab output:

```
The regression equation is
Y = 179 + 2.11 X1 + 3.56 X2 - 22.2 X3

Predictor        Coef     SE Coef          T        P
Constant      178.521       8.318      21.46    0.000
X1             2.1055      0.6479       3.25    0.008
X2            3.56240     0.09945      35.82    0.000
X3           -22.1880      0.5286     -41.98    0.000

S = 4.980       R-Sq = 99.7%     R-Sq(adj) = 99.6%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         3       89274       29758   1200.14    0.000
Residual Error    11         273          25
Total             14       89547
```

(e) 95% prediction interval for sales when $x_1 = 3, x_2 = 45, x_3 = 10$: (111.46, 135.08)

4.14

(a)

$$X'X = \begin{bmatrix} 13.00 & 59.43 & 81.82 & 115.40 \\ 59.43 & 394.73 & 360.66 & 522.08 \\ 81.82 & 360.66 & 576.73 & 728.31 \\ 115.40 & 522.08 & 728.31 & 1035.96 \end{bmatrix}$$

$$(X'X)^{-1} = \begin{bmatrix} 8.06479464 & -0.082592705 & -0.094195115 & -0.790526876 \\ -0.08259271 & 0.008479816 & 0.001716687 & 0.003720020 \\ -0.09419511 & 0.001716687 & 0.016629424 & -0.002063308 \\ -0.79052688 & 0.003720020 & -0.002063308 & 0.088601286 \end{bmatrix}$$

$$X'\boldsymbol{y} = \begin{bmatrix} 377.700 \\ 1877.911 \\ 2247.285 \\ 3339.300 \end{bmatrix}$$

(c) Estimated equation: $\hat{\mu} = 39.482 + 1.0092x_1 - 1.873x_2 - 0.367x_3$

(d) (i) (22.802, 25.653); 90% confidence interval for the mean value of y when $x_1 = 3$, $x_2 = 8$ and $x_2 = 9$ can be obtained with the software R directly using the function "predict".

Mean value	Lower limit	Upper limit
24.22764	22.80225	25.65302

There is also an option in Minitab.

(ii) (20.109, 28.346); 90% prediction interval for an individual value of y when $x_1 = 3$, $x_2 = 8$ and $x_2 = 9$ can also be obtained from the software R directly using the function "predict".

(e) F-statistic = 30.08; p-value = 0.000; reject $\beta_1 = \beta_2 = \beta_3 = 0$.

4.15

(a) Linear relationship between y and x_1; perhaps some curvature in the scatter plot of y against x_2 (see part (e))

(b) $\hat{\beta}_0 = 2.59; \hat{\beta}_1 = -0.378; \hat{\beta}_2 = 0.877$

Fitted equation: $\hat{\mu} = 2.59 - 0.378x_1 + 0.877x_2$

(c) Significant relationship between y and the variables x_1 and x_2

```
The regression equation is
Y = 2.59 - 0.378 X1 + 0.877 X2
```

Predictor	Coef	SE Coef	T	P
Constant	2.58810	0.08349	31.00	0.000
X1	-0.37802	0.06630	-5.70	0.000
X2	0.8768	0.1723	5.09	0.000

S = 0.06263   R-Sq = 90.8%   R-Sq(adj) = 89.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	0.46419	0.23210	59.16	0.000
Residual Error	12	0.04707	0.00392		
Total	14	0.51127			

(d) Model with x_1: $R^2 = 0.709$. Model with x_2: $R^2 = 0.659$. Prefer model with x_2

(e) Prefer model with both x_1 and x_2 as neither variable can be omitted from the model (see t-ratios in (c)).
No need to add $(x_2)^2$ to the model; t-ratio = 0.24; p-value = 0.815

4.16

$$(X'X)^{-1} = \begin{bmatrix} 9.61093203 & 0.008587789 & -0.27914754 & -0.04452169 \\ 0.00858779 & 0.509964070 & -0.25886359 & 0.00077654 \\ -0.27914754 & -0.25886359 & 0.13949996 & 0.00073956 \\ -0.04452169 & 0.00077654 & 0.00073956 & 0.00036978 \end{bmatrix}$$

Correction Factor = $45^2/9 = 225$
SST = $y'y$ – CF = 285 – 225 = 60
SSR = $\hat{\beta}X'y$ – CF = 282.9725 –225 = 57.9725
SSE = SST - SSR = 60 – 57.9725 = 2.0275

ANOVA table:

Source	DF	SS	MS	F	P
Regression	3	57.9725	19.3242	47.66	2.129815e-05
Residual	5	2.0275	0.4055		
Total	8	60.0000			

F-statistic = 47.66; reject $\beta_1 = \beta_2 = \beta_3 = 0$
(b) Estimated equation: $\hat{\mu} = -1.16346 + 0.13527x_1 + 0.01995x_2 + 0.12195x_3$;
$s^2 = 0.4055$;
s.e.($\hat{\beta}_0$) = 1.974; s.e.($\hat{\beta}_1$) = 0.45474; s.e.($\hat{\beta}_2$) = 0.23784; s.e.($\hat{\beta}_3$) = 0.01225
$t(\hat{\beta}_1) = 0.295$; p-value = 0.78; can not reject $\beta_1 = 0$

$t(\hat{\beta}_2) = 0.084$; p-value = 0.94; can not reject $\beta_2 = 0$
$t(\hat{\beta}_3) = 9.955$; p-value = 0.000; reject $\beta_3 = 0$

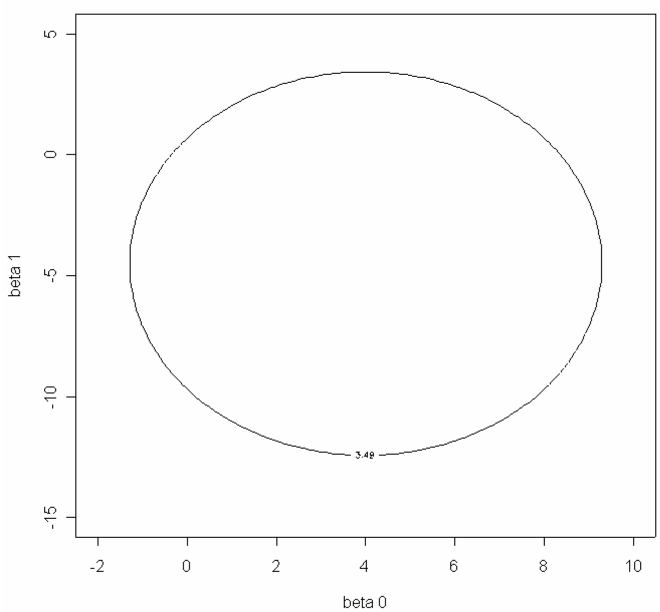
4.17

(a) $t(0.975;20) = 2.086$

95% confidence interval for $\beta_0 : 4 \pm (2.086)(2)$ or (-0.17, 8.17); covers $\beta_0 = 0$, but just barely.

95% confidence interval for $\beta_1 : -4.5 \pm (2.086)(3)$ or (-10.76, 1.76); covers $\beta_1 = 0$

(b) 95% confidence region is given below. The point $(\beta_0 = 0, \beta_1 = 0)$ is very close to the 95% contour (it is just barely within the 95% confidence region). This indicates that neither model A and B are particularly worthwhile.



(c) There is no conflict between the results in (a) and (b). In general there could have been a conflict if $\text{Cov}(\hat{\beta}_0, \hat{\beta}_1)$ was not zero.

4.18 $\text{Cov}(\boldsymbol{e}, \hat{\boldsymbol{\mu}}) = \text{Cov}((\text{I}-\text{H})\boldsymbol{y}, \text{H}\boldsymbol{y}) = \sigma^2(\text{I}-\text{H})\text{H} = \text{O}$, a (n x n) matrix of zeros. Vectors $\boldsymbol{e}$ and $\hat{\boldsymbol{\mu}}$ are linear functions of $\boldsymbol{y}$ and are normal. Hence $\boldsymbol{e}$ and $\hat{\boldsymbol{\mu}}$ are statistically independent.

4.19

(a) $\hat{\beta}^{\text{WLS}} = \sum(\text{y}_\text{i}\text{x}_\text{i}/\text{x}_\text{i}^2) \big/ \sum(\text{x}_\text{i}^2/\text{x}_\text{i}^2) = \sum(\text{y}_\text{i}/\text{x}_\text{i}) \big/ \text{n}$; $\text{V}(\hat{\beta}^{\text{WLS}}) = \sigma^2 \big/ \text{n}$

(b) $\hat{\beta}^{\text{WLS}} = 30/12 = 2.5$; $\text{V}(\hat{\beta}^{\text{WLS}}) = \sigma^2/12$

4.20

(a) $\hat{\beta}^{\text{WLS}} = \sum \text{y}_\text{i} \big/ \sum \text{x}_\text{i}$; $\text{V}(\hat{\beta}^{\text{WLS}}) = \sigma^2 \big/ \sum \text{x}_\text{i}$

(b) $\hat{\beta}^{\text{WLS}} = 30/2 = 15$; $\text{V}(\hat{\beta}^{\text{WLS}}) = \sigma^2/150$

4.21 See Exercise 4.9. Minitab output:

```
The regression equation is
y = 885 - 6.57 x1 - 1.37 x2

Predictor        Coef     SE Coef          T        P
Constant       885.16       61.75      14.33    0.000
x1            -6.5708      0.5832     -11.27    0.000
x2            -1.3743      0.1943      -7.07    0.000

S = 36.49       R-Sq = 84.0%     R-Sq(adj) = 82.8%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         2      189062       94531     71.00    0.000
Residual Error    27       35950        1331
Total             29      225011
```

95 percent confidence interval for the mean abrasion loss for rubber with hardness 70 and tensile strength 200: (134.65, 166.03)

4.22

(a) Linear model not appropriate.

(b) Fitted equation:

TensileStrength = -6.674 + 11.764 Hardwood - 0.635 (Hardwood)^2

Source	DF	SS	MS	F	P
Regression	2	3104.2	1552.1	79.43	0.000
Residual Error	16	312.6	19.5		
Total	18	3416.9			

Model adequate; quadratic term needed; increases R^2 from 0.305 to 0.909.
95% confidence interval for mean response when hardwood 6 percent: (38.14, 44.00)
Prediction intervals are for individual observations while confidence intervals are for the mean value. Confidence intervals are shorter than the corresponding prediction intervals. 95% prediction interval for tensile strength for a batch of paper with 6 percent hardwood concentration: (31.25, 50.88).
The maximum hardwood concentration in the data set used to fit the model is 7 percent, which is very low compared to 20 percent. It is not advisable to use the fitted model to predict the mean tensile strength of paper for 20 percent hardwood concentration.

4.23

Quadratic model. Estimated equation: $\hat{\mu} = 82.385 - 38.310\text{x} + 4.703\text{x}^2$
Regression significant; adequate fit.
Stars with ln(surface temperature) < 4 appear different and should be investigated separately. Without these stars, a linear model is appropriate.

Predictor	Coef	SE Coef	T	P
Constant	82.385	9.581	8.60	0.000
x	-38.310	4.790	-8.00	0.000
x2	4.7025	0.5939	7.92	0.000

```
S = 0.3667      R-Sq = 60.6%     R-Sq(adj) = 58.8%

Analysis of Variance
```

Source	DF	SS	MS	F	P
Regression	2	9.0945	4.5472	33.82	0.000
Residual Error	44	5.9165	0.1345		
Total	46	15.01			

4.24

(a) Minitab regression output

```
The regression equation is
Y = 31.4 + 9.31 UFFI + 2.85 Tight
```

Predictor	Coef	SE Coef	T	P
Constant	31.373	2.461	12.75	0.000
UFFI	9.312	2.133	4.37	0.000
Tight	2.8545	0.3764	7.58	0.000

```
S = 5.223       R-Sq = 78.3%     R-Sq(adj) = 76.2%
```

```
Analysis of Variance

Source            DF        SS        MS        F        P
Regression         2    2063.3    1031.6    37.82    0.000
Residual Error    21     572.9      27.3
Total             23    2636.1
```

(b)

$$(X'X) = \begin{bmatrix} 24 & 12 & 123 \\ 12 & 12 & 61 \\ 123 & 61 & 823 \end{bmatrix}; \ (X'X)^{-1} = \begin{bmatrix} 0.221946 & -0.085569 & -0.026828 \\ -0.085569 & 0.166703 & 0.000433 \\ -0.026828 & 0.000433 & 0.005193 \end{bmatrix};$$

$$V(\hat{\boldsymbol{\beta}}) = s^2 (X'X)^{-1} = \begin{bmatrix} 6.05462 & -2.33430 & -0.73187 \\ -2.33430 & 4.54761 & 0.01180 \\ -0.73187 & 0.01180 & 0.14165 \end{bmatrix}$$

$\text{s.e.}(\hat{\beta}_0) = \sqrt{6.05462} = 2.461$; $\text{s.e.}(\hat{\beta}_1) = \sqrt{4.54761} = 2.133$;
$\text{s.e.}(\hat{\beta}_2) = \sqrt{0.14165} = 0.376$

(c) The 95 percent confidence region for (β_1, β_2) is shown below. The point $(\beta_1 = 0, \beta_2 = 0)$ is far from this region.