SOLUTIONS TO EXERCISES IN CHAPTER 2

2.1 (a) We use the relationship $\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}(\mathbf{X'Y})$. The vector $\mathbf{X'Y}$ is of the form

$$\mathbf{X'Y} = \begin{bmatrix} \Sigma y_t \\ \Sigma x_{t1} y_t \\ \Sigma x_{t2} y_t \end{bmatrix} = \begin{bmatrix} 10 \\ 40 \\ 40 \end{bmatrix},$$ and the matrix $\mathbf{X'X}$ is of the form

$$\mathbf{X'X} = \begin{bmatrix} n & \Sigma x_{t1} & \Sigma x_{t2} \\ \Sigma x_{t1} & \Sigma x_{t1}^2 & \Sigma x_{t1} x_{t2} \\ \Sigma x_{t2} & x_{t1} x_{t2} & \Sigma x_{t2}^2 \end{bmatrix} = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 40 \end{bmatrix},$$

so that $$(\mathbf{X'X})^{-1} = \begin{bmatrix} 1/10 & 0 & 0 \\ 0 & 1/20 & 0 \\ 0 & 0 & 1/40 \end{bmatrix} = \begin{bmatrix} .10 & 0 & 0 \\ 0 & .05 & 0 \\ 0 & 0 & .025 \end{bmatrix}$$

and then, $$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} .10 & 0 & 0 \\ 0 & .05 & 0 \\ 0 & 0 & .025 \end{bmatrix} \cdot \begin{bmatrix} 10 \\ 40 \\ 40 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

(b) We calculate $\hat{\boldsymbol{\beta}}'\mathbf{X'Y} = [\,1 \quad 2 \quad 1\,] \cdot \begin{bmatrix} 10 \\ 40 \\ 40 \end{bmatrix} = 130$, and $\overline{y} = \frac{1}{n} \cdot \Sigma y_t = \frac{1}{10} \cdot 10 = 1$,

$\mathbf{y'y} = \Sigma y_t^2 = 165$, so that
SSTO $= \mathbf{y'y} - n\overline{y}^2 = 165 - 10(1^2) = 155$, and SSE $=$ SSTO $-$ SSR $= 35$

The ANOVA table is

Source	Sum of Squares	degrees of freedom	Mean Square	F-Ratio
Regression	SSR $= \hat{\boldsymbol{\beta}}'\mathbf{X'Y} - n\overline{y}^2$ $= 120$	$p = 2$	MSR $= \frac{\text{SSR}}{p}$ $= 60$	$\frac{\text{MSR}}{\text{MSE}} = 12$
Residual	SSE $= \mathbf{e'e}$ $= 35$	$n - p - 1$ $= 7$	MSE $= \frac{\text{SSE}}{n-p-1}$ $= 5$	

Total SSTO $= \mathbf{y'y} - n\bar{y}^2 = 155$ 9

$R^2 = \frac{SSR}{SSTO} = \frac{120}{155} = .774$, which indicates a fairly strong relationship between y and x_1,x_2.

(c) The standard error of $\widehat{\beta}_0$ is $s_{\widehat{\beta}_0} = s\sqrt{c_{00}}$, where $s = \sqrt{MSE} = \sqrt{5}$ and c_{ij} is the i, j-entry in the matrix $\mathbf{C = (X'X)^{-1}}$, so that $c_{00} = .10$ ($\mathbf{(X'X)^{-1}}$ is a $(p+1)$x$(p+1)$ matrix, and the rows and columns are numbered from 0 to p). Thus, $s_{\widehat{\beta}_0} = \sqrt{5} \cdot \sqrt{.1} = .7071$.
In a similar way, we have $s_{\widehat{\beta}_1} = s\sqrt{c_{11}} = .5$, and $s_{\widehat{\beta}_2} = s\sqrt{c_{22}} = .3536$.

2.1 (c) The t-statistics for $\widehat{\beta}_0$, $\widehat{\beta}_1$ and $\widehat{\beta}_2$ are $t_{\widehat{\beta}_0} = \frac{\widehat{\beta}_0}{s_{\widehat{\beta}_0}} = \frac{1}{.7071} = 1.414$,

$t_{\widehat{\beta}_1} = \frac{\widehat{\beta}_1}{s_{\widehat{\beta}_1}} = \frac{2}{.5} = 4.0$ and $t_{\widehat{\beta}_2} = \frac{\widehat{\beta}_2}{s_{\widehat{\beta}_2}} = \frac{1}{.3536} = 2.828$.

The residual degrees of freedom are 7, so that $t_{.025}(7) = 2.36$.
According to the calculated t-statistics for $\widehat{\beta}_0$, $\widehat{\beta}_1$ and $\widehat{\beta}_2$, for the hypothesis test with null hypothesis H_0: $\beta_i = 0$ and alternative H_1: $\beta_i \neq 0$, at significance level $\alpha = .05$, we would not reject H_0 in the test of β_0 (since $1.414 = |t| \leq t_{.025}(7) = 2.36$) , but we would reject H_0 in the cases of β_1 (since $4.0 > 2.36$) and β_2 (since $2.828 > 2.36$).

(d) The F-ratio for the hypothesis test with null hypothesis H_0: $\beta_1 = \beta_2 = 0$ and alternative H_1: *at least one of* β_1 *or* $\beta_2 \neq 0$ has F-ratio (from the ANOVA table) $F = 12$. The critical value for the test with a level of significance of $\alpha = .05$ is $F_\alpha(p, n-p-1) = F_{.05}(2,7) = 4.74$.
Thus, H_0 is rejected at the 5% level of significance (since $12 > 4.74$).

2.2 In this simple linear regression exercise, $x_t = t$, for $t = 1, ..., n$. The normal equations are:

$n\beta_0 + \beta_1\sum_1^n x_t = \sum_1^n y_t \rightarrow n\beta_0 + \beta_1\sum_1^n t = \sum_1^n y_t \rightarrow n\beta_0 + \frac{n(n+1)}{2}\beta_1 = \sum_1^n y_t$ and

$\beta_0\sum_1^n x_t + \beta_1\sum_1^n x_t^2 = \sum_1^n x_t y_t \rightarrow \beta_0\sum_1^n t + \beta_1\sum_1^n t^2 = \sum_1^n t y_t \rightarrow \frac{n(n+1)}{2}\beta_0 + \frac{n(n+1)(2n+1)}{6}\beta_1 = \sum_1^n t y_t.$

Since the x_t's are the integers from 1 to n, it follows that $\overline{x} = \frac{n+1}{2}$, and

$n\Sigma(x_t - \overline{x})^2 = n\Sigma x_t^2 - (\Sigma x_t)^2 = n\frac{n(n+1)(2n+1)}{6} - \left(\frac{n(n+1)}{2}\right)^2 = \frac{n^2(n^2-1)}{12}$, so that

the least squares estimates are:

$$\widehat{\beta}_1 = \frac{n\Sigma x_t y_t - (\Sigma x_t)(\Sigma y_t)}{n\Sigma x_t^2 - (\Sigma x_t)^2} = \frac{n\Sigma t y_t - \frac{n(n+1)}{2}\Sigma y_t}{\frac{n^2(n^2-1)}{12}} = \frac{12\Sigma t y_t - 6(n+1)\Sigma y_t}{n(n-1)(n+1)}$$

and $\widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1\overline{x} = \overline{y} - \frac{12\Sigma t y_t - 6(n+1)\Sigma y_t}{n(n-1)(n+1)} \cdot \frac{n+1}{2} = \overline{y} - \frac{6\Sigma t y_t - 3(n+1)\Sigma y_t}{n(n-1)}$.

2.3 (a) The t-value used to test $H_0 : \beta_1 = 0$ vs. $H_1 \neq 0$ is $t_{\widehat{\beta}_1} = \widehat{\beta}_1/s_{\widehat{\beta}_1} = \frac{1.14}{.16} = 7.25$.
With $n = 17$ and $p = 2$, for a test at level $\alpha = .05$, we find from the t-tables that $t_{.025}(n-p-1) = t_{.025}(14) = 2.14$. Since $\left|t_{\widehat{\beta}_1}\right| > t_{.025}(n-p-1)$, we reject H_0 at the 5% level.

(b) The completed ANOVA table is

Source	SS	df	MS
Regression	66	2	33
Error	34	14	2.429
Total (corr. for mean)	100	16	

The entries in the table are found as follows: SSR = SSTO − SSE = 100 − 34 = 66;
the regression degrees of freedom are $p = 2$, the error df are $n - p - 1 = 17 - 2 - 1 = 14$
and the total df are $n - 1 = 16$; MSR = SSR/p = 66/2 = 33,
MSE = SSE/$(n - p - 1)$ = 34/14 = 2.429.
R^2 = SSR/SSTO = 66/100 = .66.
To test $H_0 : \beta_1 = \beta_2 = 0$ we calculate $F = \frac{MSR}{MSE} = \frac{33}{2.429} = 13.59$. For a test at the

$\alpha = .05$ level, we find $F_\alpha(p, n-p-1) = F_{.05}(2, 14) = 3.74$. Since $F > F_{.05}$, we reject H_0 at the 5% level.

2.4 (a) The design matrix is $\mathbf{X} = \begin{bmatrix} 1 & cos\theta & sin\theta \\ 1 & -sin\theta & -cos\theta \\ 1 & sin\theta & cos\theta \\ 1 & -cos\theta & -sin\theta \end{bmatrix}$

Then $\mathbf{X'X} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 2sin2\theta \\ 0 & 2sin2\theta & 2 \end{bmatrix}$ and $\mathbf{X'Y} = \begin{bmatrix} y_1 + y_2 + y_3 + y_4 \\ (y_1 - y_4)cos\theta + (y_3 - y_2)sin\theta \\ (y_1 - y_4)sin\theta + (y_3 - y_2)cos\theta \end{bmatrix}$

and $\mathbf{(X'X)^{-1}} = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{2cos^2 2\theta} & \frac{-sin2\theta}{2cos^2 2\theta} \\ 0 & \frac{-sin2\theta}{2cos^2 2\theta} & \frac{1}{2cos^2 2\theta} \end{bmatrix}$ so that

$\boldsymbol{\widehat{\beta}} = \mathbf{(X'X)^{-1}X'Y}$, which results in $\widehat{\beta}_0 = (y_1 + y_2 + y_3 + y_4)/4$,

$\widehat{\beta}_1 = \frac{(y_1-y_4)[cos\theta - sin\theta\, sin2\theta] + (y_3-y_2)[sin\theta - cos\theta\, sin2\theta]}{2cos^2 2\theta} = \frac{(y_1-y_4)cos\theta - (y_3-y_2)sin\theta}{2cos2\theta}$

since $cos\theta - sin\theta\, sin2\theta = cos\theta - sin\theta\, 2sin\theta\, cos\theta = cos\theta[1-2sin^2\theta] = cos\theta\, cos2\theta$, and $sin\theta - cos\theta\, sin2\theta = sin\theta - cos\theta\, 2sin\theta\, cos\theta = sin\theta\, cos2\theta$, and similarly,

$\widehat{\beta}_2 = \frac{(y_1-y_4)[sin\theta - cos\theta\, sin2\theta] + (y_3-y_2)[cos\theta - sin\theta\, sin2\theta]}{2cos^2 2\theta} = \frac{-(y_1-y_4)sin\theta + (y_3-y_2)cos\theta}{2cos2\theta}$

(b) $V(\widehat{\beta}_1) = \sigma^2 c_{11} = \frac{\sigma^2}{2cos^2\theta}$ (c_{11} is the (1,1)-entry in $\mathbf{(X'X)^{-1}}$), and similarly $V(\widehat{\beta}_2) = \sigma^2 c_{22} = \frac{\sigma^2}{2cos^2\theta}$.

2.5 (a) The ANOVA table is

Source	SS	df	MS
Regression	66	2	33
Error	134	17	7.882
Total (corr. for mean)	200	19	

The entries in the ANOVA table are calculated as follows:
SSE = SSTO − SSR = 200 − 66 = 134; regression df are $p = 2$, error df are $n - p - 1 = 17$, total df are $n - 1 = 19$; MSE = SSE/p = 66/2 = 33,
MSR = SSR/$(n - p - 1)$ = 134/17 = 7.882 .

(b) $R^2 = \frac{\text{SSR}}{\text{SSTO}} = \frac{66}{200} = .33$

(c) The F-ratio is $F = \frac{\text{MSR}}{\text{MSE}} = \frac{33}{7.882} = 4.187$,
and the critical value for $\alpha = .05$ is $F_{.05}(2, 17) = 3.59$, so that $H_0: \beta_1 = \beta_2 = 0$ is rejected at the .05 significance level, since $F > F_{.05}(p, n-p-1)$.

2.5 (d) We can test whether $\beta_2 = 0$ by using the extra sum of squares approach. From the original information, we know that $\text{SSR}(X_1, X_2) = 66$, and we are now told that $\text{SSR}(X_1) = 50$ (for the simpler one-variable regression model considered in this part of the problem). The extra regression sum of squares is

$$\begin{aligned}\text{SSR}(X_{q+1}, ..., X_p \mid X_1, X_2, ..., X_q) &= \text{SSR}(X_1, X_2, ..., X_q, ..., X_p) - \text{SSR}(X_1, X_2, ..., X_q) \\ &= \text{SSR}(X_1, X_2) - \text{SSR}(X_1) = 16\end{aligned}$$

(here $p = 2$ and $q = 1$). Then the F statistic to test $H_0: \beta_{q+1} = \cdots = \beta_p = 0$,

which in this case is $H_0: \beta_2 = 0$, is $F^* = \frac{\text{SSR}(X_{q+1},...,X_p|X_1,X_2,...,X_q)/(p-q)}{\text{SSE}(X_1,X_2,...,X_p)/(n-p-1)}$
so that $F^* = \frac{16/1}{134/17} = 2.030$.
The critical value for the hypothesis test at level $= .05$ is $F_{.05}(1, 17) = 4.45$, so that H_0 is not rejected at the 5% level of significance, since $2.030 < 4.45$.

2.6 (a) $R^2 = \frac{\text{SSR}}{\text{SSTO}} \rightarrow \text{SSR} = (.88)(100) = 88$, and then $\text{SSE} = \text{SSTO} - \text{SSR} = 100 - 88 = 12$.
The F-ratio to test $H_0: \beta_1 = \beta_2 = \beta_3 = 0$ is $F = \frac{\text{SSR}/p}{\text{SSE}/(n-p-1)} = \frac{88/3}{12/6} = 14.67$, and
$F_{.05}(3, 6) = 4.76$, so that H_0 is rejected at the 5% level.

(b) From the information given, we can find $\text{SSR}(X_1, X_2) = \text{SSR}(X_1 \mid X_2) + \text{SSR}(X_2) = 85$,
$\text{SSR}(X_2, X_3) = \text{SSR}(X_3 \mid X_2) + \text{SSR}(X_2) = 41$,
$\text{SSR}(X_1, X_3) = \text{SSR}(X_1, X_2, X_3) - \text{SSR}(X_2 \mid X_1, X_3) = 88 - 2 = 86$
(note that $\text{SSR}(X_1, X_2, X_3)$ is SSR for the full 3-variable regression from part (a)).
The partial test for X_1 (with $\alpha = .05$) uses F-ratio

$$F^* = \frac{\text{SSR}(X_1|X_2,X_3)/1}{\text{SSE}(X_1,X_2,X_3)/(n-p-1)} = \frac{\text{SSR}(X_1,X_2,X_3) - \text{SSR}(X_2,X_3)}{\text{SSE}(X_1,X_2,X_3)/6} = \frac{88-41}{12/6} = 23.5,$$

and critical value $F_{.05}(1, 6) = 5.99$, so that X_1 appears to make a significant additional effect on Y after X_2 and X_3 have already been included.

The partial test for X_2 (with $\alpha = .05$) uses F-ratio

$$F^* = \frac{\text{SSR}(X_2|X_1,X_3)/1}{\text{SSE}(X_1,X_2,X_3)/(n-p-1)} = \frac{\text{SSR}(X_1,X_2,X_3) - \text{SSR}(X_1,X_3)}{\text{SSE}(X_1,X_2,X_3)/6} = \frac{88-86}{12/6} = 1.0,$$

and critical value $F_{.05}(1, 6) = 5.99$, so that X_2 appears to make no significant additional effect on Y after X_1 and X_3 have already been included.

The partial test for X_3 (with $\alpha = .05$) uses F-ratio

$$F^* = \frac{\text{SSR}(X_3|X_1,X_2)/1}{\text{SSE}(X_1,X_2,X_3)/(n-p-1)} = \frac{\text{SSR}(X_1,X_2,X_3) - \text{SSR}(X_1,X_2)}{\text{SSE}(X_1,X_2,X_3)/6} = \frac{88-85}{12/6} = 1.5,$$

and critical value $F_{.05}(1, 6) = 5.99$, so that X_3 appears to make no significant additional effect on Y after X_1 and X_2 have already been included.

(c) To test $H_0: \beta_2 = \beta_3 = 0$, we use the extra sum of squares approach and find

$$F^* = \frac{\text{SSR}(X_2,X_3|X_1)/(3-1)}{SSE(X_1,X_2,X_3)/(10-3-1)} = \frac{[\text{SSR}(X_1,X_2,X_3) - \text{SSR}(X_1)]/2}{\text{SSE}(X_1,X_2,X_3)/6} = \frac{(88-82)/2}{12/6} = 1.5,$$

and the critical value is $F_{.05}(2,6) = 5.14$. Thus, we do not reject H_0 at the .05 level.

(d) These tests clearly point to X_1 as being the most important variable in describing the effect on Y, with X_2 and X_3 being of little or no significance.

2.7 (a) There are $3+3+3+3+5 = 17 = n$ data points and $k = 5$ settings of X.
The ANOVA table is

Source	SS	df	MS	F
Regression	73	1	73	36.5
Error	30	15	2.0	
Lack of fit	21	3	7.0	
Pure error	9	12	0.75	
Total (corr. for mean)	103	16		

The entries in the ANOVA table are found as follows: MSR $=$ SSR/1 $= 73$,
SSE has $n-p-1 = 17-1-1 = 15$ df and MSE $=$ SSE/15 $= 30/15 = 2.0$,
$F =$ MSR/MSE $= 73/2.0 = 36.5$; SSLF $=$ SSE $-$ SSPE $= 30-9 = 21$,
SSLF has $k-p-1 = 5-1-1 = 3$ df , MSLF $=$ SSLF/3 $= 21/3 = 7.0$,
SSPE has $n-k = 17-5 = 12$ df , MSPE $=$ SSPE/12 $= 9/12 = 0.75$.
SSTO has $n-1 = 16$ df.

(b) We can test $H_0 : \beta_1 = 0$ using the usual F-test: $F = 36.5$ and $F_{.05}(1,15) = 4.54$,
so that H_0 is rejected at the 5% level - there appears to be a significant effect on Y by X.

We can perform a lack of fit F-test: $F_{\text{LF}} = \frac{\text{SSLF}/(k-p-1)}{\text{SSPE}/(n-k)} = \frac{\text{MSLF}}{\text{MSPE}} = \frac{7.0}{0.75} = 9.33$,
and since $F_{.05}(3,12) = 3.49$, we conclude that lack of fit is indicated in this model.

(c) Usually $s^2 =$ MSR (2.0 in this example) is taken as the estimate of σ^2. Since the model appears to suffer from lack of fit, the appropriate estimate of σ^2 is MSPE $= 0.75$.

2.8 (a) The model can be written as $y_t = \beta x_t + \epsilon_t$, where $y_t = z_t - 100$ and $x_t = y_{t-1}$.
We have 9 data points (x_2, y_2) , . . . , (x_{10}, y_{10}) (there is no x_1, since this would be y_0).
This is now a simple regression model through the origin, so that
$\widehat{\beta} = \frac{\Sigma x_t y_t}{\Sigma x_t^2} = \frac{x_2y_2+\cdots+x_{10}y_{10}}{x_2^2+\cdots+x_{10}^2} = \frac{y_1y_2+\cdots+y_9y_{10}}{y_1^2+\cdots+y_9^2} = \frac{160}{216-16} = .80$

(note that $y_1^2 + \cdots + y_9^2 = y_1^2 + \cdots + y_9^2 + y_{10}^2 - y_{10}^2$).

(b) $\widehat{y}_{11} = \widehat{\beta}x_{11} = \widehat{\beta}y_{10} =$ (.80)(4) $= 3.2 \;\rightarrow\; \widehat{z}_{11} = \widehat{y}_{11} + 100 = 103.2$,
and $\widehat{y}_{12} = \widehat{\beta}y_{11} \doteq \widehat{\beta}\widehat{y}_{11} =$ (.80)(3.2) $= 2.56 \;\rightarrow\; \widehat{z}_{12} = 102.56$.

(c) The 95% prediction interval for z_{11} is $\widehat{z}_{11} \pm t_{.025}(n-p)\,\text{s}\sqrt{1 + \frac{x_{11}^2}{\Sigma x_t^2}}$.
$x_{11} = y_{10}$, and $\sum_{2}^{10} x_t^2 = \sum_{1}^{9} y_t^2$, and the df is $n-p$ because
there is no constant term in the regression model. The interval becomes

$$\widehat{z}_{11} \pm t_{.025}(8) \cdot 1 \cdot \sqrt{1 + \frac{y_{10}^2}{\sum_{1}^{9} y_t^2}} = 103.2 \pm (2.31)\sqrt{1 + \frac{16}{200}} = 103.2 \pm 2.40 \,.$$

2.9 The wording suggests that $s =$ (.7)$\sqrt{\text{SSTO}/(n-1)}$, since without the regression, the residual standard deviation is the sample standard deviation of the y_t's , i.e., $\sqrt{\text{SSTO}/n-1}$, whereas, with the regression, the residual standard deviation is $s = \sqrt{\text{SSE}/(n-p-1)}$. But then
$s^2 =$ (.49)[SSTO/$(n-1)$] $\;\rightarrow\; \frac{\text{SSE}/(n-p-1)}{\text{SSTO}/(n-1)} = .49 \;\rightarrow\; R_a^2 = 1 - \frac{\text{SSE}/(n-p-1)}{\text{SSTO}/(n-1)} = .51$.

2.10 (a) $\widehat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{y}$, and $\widehat{\mathbf{y}}' = \boldsymbol{\beta}'\mathbf{X}'$ so that
$\widehat{\mathbf{y}}'\mathbf{y} = \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{y} = \boldsymbol{\beta}'\mathbf{X}'\mathbf{y}$ (since the $\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \mathbf{I}$).

(b) $F = \frac{\text{MSR}}{\text{MSE}} = \frac{\text{SSR}/p}{\text{SSE}/(n-p-1)} = \frac{\text{SSR}(n-p-1)}{(\text{SSTO}-\text{SSR})p} = \frac{\frac{\text{SSR}}{\text{SST0}}(n-p-1)}{(\frac{\text{SSTO}}{\text{SSTO}}-\frac{\text{SSR}}{\text{SSTO}})p} = \frac{R^2}{1-R^2}\ \frac{n-p-1}{p}$

2.11 If estimation had been done using the true model $y_t = \beta_0 + \beta_1 x_t + \epsilon_t$, then in the course of the estimation, it is found that $\widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1\overline{x}$ is an unbiased estimate of β_0. Thus, using $\overline{y}$ as an estimate of β_0 results in a bias of $\widehat{\beta}_1\overline{x}$.

2.12 (a) Since $E(\mathbf{y}) = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2$, it follows that
$\mathrm{E}(\widehat{\boldsymbol{\beta}}_1) = \mathrm{E}[(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y}] = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'[\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2] = \boldsymbol{\beta}_1 + (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2 \boldsymbol{\beta}_2$

(b) If the $\mathbf{X_1}$ is "orthogonal" to $\mathbf{X_2}$, i.e., $\sum_{t=1}^{n} x_{ti}x_{tj} = 0$ for any x_i from $\mathbf{X_1}$ and x_j from $\mathbf{X_2}$,

then $\mathbf{X}_1'\mathbf{X_2} = \mathbf{0}$ (a matrix with all entries 0), and the $\widehat{\boldsymbol{\beta}}_1$ from part (a) will be unbiased.

2.13 (a) $V(\widehat{\beta}_1 - \widehat{\beta}_2) = V(\widehat{\beta}_1) + V(\widehat{\beta}_2) = \frac{\sigma_1^2}{\Sigma(x_{t1}-\overline{x}_1)^2} + \frac{\sigma_2^2}{\Sigma(x_{t2}-\overline{x}_2)^2}$, since $\widehat{\beta}_1$ and $\widehat{\beta}_2$ are

independent (they are found from independent samples). σ_1^2 and σ_2^2 are estimated as
$s_1^2 = \text{SSE}_1/(n-p-1) = 21{,}357/13 = 1642.85$ and
$s_2^2 = \text{SSE}_2/(n-p-1) = 29{,}064/13 = 2235.69$. The estimated variance is
$\widehat{V}(\widehat{\beta}_1 - \widehat{\beta}_2) = \widehat{V}(\widehat{\beta}_1) + \widehat{V}(\widehat{\beta}_2) = \frac{s_1^2}{\Sigma(x_{t1}-\overline{x}_1)^2} + \frac{s_2^2}{\Sigma(x_{t2}-\overline{x}_2)^2} = \frac{1642.85}{15{,}240} + \frac{2235.69}{15{,}001} = .257$.

To test $H_0 : \beta_1 = \beta_2$, we calculate the t-value $t = \frac{\widehat{\beta}_1 - \widehat{\beta}_2}{\sqrt{\widehat{V}(\widehat{\beta}_1 - \widehat{\beta}_2)}} = \frac{2.21-3.59}{\sqrt{.257}} = -2.72$.

The critical value for the hypothesis test with $\alpha = .10$ is $t_{.05}(26) = 1.71$ (df are $13 + 13$ from the two samples). Since $|t| > t_{.05}(26)$, H_0 is rejected at the 10% level of significance.

(b) $V(\widehat{\beta}_1 - \widehat{\beta}_2) = V(\widehat{\beta}_1) + V(\widehat{\beta}_2)$ is still true.

2.14 (a) Combining the data results in an n x 3 design matrix $\mathbf{X} = \begin{bmatrix} 1 & -1 & -c \\ \vdots & \vdots & \vdots \\ 1 & -1 & -c \\ 1 & -1 & c \\ \vdots & \vdots & \vdots \\ 1 & -1 & c \\ 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 0 \end{bmatrix}$.

The model can be written as $\mathbf{Y}, \mathbf{Z} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$,
indicating that both Y and Z are incorporated into the

model.

2.14 (b) The normal equations are $(\mathbf{X'X})\widehat{\boldsymbol{\beta}} = \mathbf{X'y}$, which in this case becomes

$$\begin{bmatrix} n & n-4m & 0 \\ n-4m & n & 0 \\ 0 & 0 & 2mc^2 \end{bmatrix} \begin{bmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix} = \begin{bmatrix} \Sigma y_t \\ -\sum_{1}^{2m} y_t + \sum_{2m+1}^{n} y_t \\ -c\sum_{1}^{m} y_t + c\sum_{m+1}^{2m} y_t \end{bmatrix}$$

$$\mathbf{V}(\widehat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X'X})^{-1} = \sigma^2 \begin{bmatrix} \frac{n}{8mn-16m^2} & \frac{-n+4m}{8mn-16m^2} & 0 \\ \frac{-n+4m}{8mn-16m^2} & \frac{n}{8mn-16m^2} & 0 \\ 0 & 0 & \frac{1}{2mc^2} \end{bmatrix}$$

The variance of $\widehat{y}$ (as a prediction) is $V(\widehat{y}) = \sigma^2[\,1 + \mathbf{x}_k'(\mathbf{X'X})^{-1}\mathbf{x}_k\,]$.
In this case, $\mathbf{x}_k' = [\,1 \;\; 1 \;\; c\,]$ and $n = 7$, so that $V(\widehat{y}) = \sigma^2[\,1 + \frac{1}{7-2m} + \frac{1}{2m}\,]$
The minimum $V(\widehat{y})$ occurs at the minimum of $g(m) = \frac{1}{7-2m} + \frac{1}{2m}$. Since $n = 7$, and $m \le \frac{m}{2}$, we must have $m \le 3$ (m is an integer). First, $m \neq 0$ since division by 0 is not allowed. By trial and error, if $m = 1$, $g(1) = .7$, if $m = 2$, $g(2) = .583$, and if $m = 3$, $g(3) = 1.17$. The minimum occurs at $m = 2$.

2.15 The regression model is $y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \epsilon_t$.
Using a computer solution for the problem, we get
$\widehat{y}_k = 2.483 + 2.292x_{k1} - .050x_{k2} + 13.924x_{k3} - .042x_{k4}$.
With $x_{k1} = 10$, $x_{k2} = 6.0$, $x_{k3} = 1.5$ and $x_4 = 20$, the predicted sale value will be $\widehat{y}_k = 45.09$ (thousands). $s^2 = \text{MSE} = 17.55$ (with $n - p - 1 = 23$ df) $\rightarrow$ $s = 4.19$.
The 95% prediction interval is $\widehat{y}_k \pm t_{.025}(23)\, s\sqrt{1 + \mathbf{x}_k'(\mathbf{X'X})^{-1}\mathbf{x}_k}$, where
$\mathbf{x}_k' = [\,1 \;\; 10 \;\; 6 \;\; 1.5 \;\; 20\,]$. This interval is $45.09 \pm (2.07)(4.19)\sqrt{1.19} = 45.09 \pm 9.46$.

2.16 (a) False. Least squares estimation introduces restrictions among residuals.

(b) True. $\text{SSR} = \widehat{\boldsymbol{\beta}}'\mathbf{X'y} - n\overline{y}^2$, but $\widehat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'y}$ so that
$\widehat{\boldsymbol{\beta}}' = [(\mathbf{X'X})^{-1}\mathbf{X'y}]' = \mathbf{y'X}\,[(\mathbf{X'X})^{-1}]' = \mathbf{y'X}\,(\mathbf{X'X})^{-1}$
($\mathbf{X'X}$ is a symmetric matrix and so is its inverse $(\mathbf{X'X})^{-1}$, which means that $(\mathbf{X'X})^{-1}$ is equal to its own transpose).

(c) True. SSE will never increase (and will usually decrease) with the addition of a variable into a regression model, but MSE $=$ SSE/$(n-p-1)$ may increase as the number of variables in model goes from p to $p+1$.

(d) False. An t-test of a single β_i is a test of significance of β_i after all other variables have been included in the model. A β_i may be insignificant because the variable X_i is strongly correlated with another variable and doesn't provide any "significant" additional description of the behavior of Y.

2.16 (e) False.

(f) False. $R^2 = \frac{\text{SSR}}{\text{SSTO}} = \frac{\Sigma(\widehat{y}_t - \overline{y})^2}{\Sigma(y_t - \overline{y})^2}$.

(g) False. The definition of s_x^2 is $s_x^2 = \frac{\Sigma(x_t - \overline{x})^2}{n-1}$, and $s_y^2 = \frac{\Sigma(y_t - \overline{y})^2}{n-1}$.

$$\widehat{\beta}_1 = \frac{\Sigma(x_t-\overline{x})(y_t-\overline{y})}{\Sigma(x_t-\overline{x})^2} \quad \text{and} \quad r = \frac{\Sigma(x_t-\overline{x})(y_t-\overline{y})}{\sqrt{\Sigma(x_t-\overline{x})^2\,\Sigma(y_t-\overline{y})^2}} \quad \longrightarrow \quad \frac{\widehat{\beta}_1}{r} = \sqrt{\frac{\Sigma(y_t-\overline{y})^2}{\Sigma(x_t-\overline{x})^2}} = \frac{s_y}{s_x}$$

(the $n-1$'s cancel).

(h) False.

(i) True. See Example 1 on p. 59 of the text.

(j) False. A qualitative variable with k levels is represented by $k-1$ indicator or "dummy variables.

(k) False. The model will have to be transformed so that the error term is additive.

(l) True. Use the transformation $z_t = -\,ln(\frac{1}{y_t} - 1) = \beta_0 + \beta_1 x_t + ln(\epsilon_t)$.

(m) False. The procedure allows for a lack-of-fit test for any number p of independent variables. See p. 57 of the text.

(n) False. As mentioned on p. 47, "... it is quite possible that backward elimination and forward selection lead to models that include different explanatory variables ...".

(o) False. Correlation is not causation.

(p) False. Residuals should always be checked for correlation.

(q) False. The sum of residuals from a linear regression is 0 if the model includes a constant term.

2.17 (a) $y_t = \beta_0 + \beta_1 t + \delta_1 \text{IND}_t + \epsilon_t$, where $\text{IND}_t = 1$ if $t \geq 0$ and is 0 otherwise. This is a 2-variable model, variable 1 is time, t, and variable 2 is the indicator variable, IND_t which is 0 or 1 depending upon whether the treatment has not yet been applied ($t < 0$) or has been applied $(t \geq 0)$. The design matrix is

$$\mathbf{X} = \begin{bmatrix} \mathbf{1} & -\mathbf{5} & \mathbf{0} \\ \mathbf{1} & -\mathbf{3} & \mathbf{0} \\ \mathbf{1} & -\mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{3} & \mathbf{1} \\ \mathbf{1} & \mathbf{5} & \mathbf{1} \end{bmatrix} \text{ and } \mathbf{X'X} = \begin{bmatrix} 6 & 0 & 3 \\ 0 & 70 & 9 \\ 3 & 9 & 3 \end{bmatrix} , \ \mathbf{X'y} = \begin{bmatrix} 71 \\ 99 \\ 49 \end{bmatrix} .$$

A computer generated solution results in parameter estimates of $\widehat{\beta}_0 = 10.708$, $\widehat{\beta}_1 = 1.125$, $\widehat{\delta}_1 = 2.250$. The t-value for the t-test of significance of δ_1 is $t_{\widehat{\delta}_1} = 2.192$, and since $t_{.025}(3) = 3.18$ ($n - p - 1 = 6 - 2 - 1 = 3$ df) , we see that $H_0 : \delta_1 = 0$ is not rejected at the 5% level of significance (however $H_0 : \beta_1 = 0$ is rejected at the 5% level). The data indicates that the jump is not significant.

2.17 (b) This situation can be modeled by $y_t = \beta_0 + \beta_1 t + \beta_2 t\,\text{IND}_t + \epsilon_t$, where IND_t is the same as in part (a). The design matrix is now

$$\mathbf{X} = \begin{bmatrix} \mathbf{1} & -\mathbf{5} & \mathbf{0} \\ \mathbf{1} & -\mathbf{3} & \mathbf{0} \\ \mathbf{1} & -\mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{3} & \mathbf{3} \\ \mathbf{1} & \mathbf{5} & \mathbf{5} \end{bmatrix}.$$

The computer generated solution results in $\widehat{\beta}_0 = 12.208$, $\widehat{\beta}_1 = 1.539$, $\widehat{\beta}_2 = -.250$, but the t-test at the 5% level of $H_0 : \beta_2 = 0$ is not rejected (it is rejected for β_1).

2.18 As mentioned on p. 51 of the text, "if the usual least squares procedures are employed in the presence of serially correlated errors, the parameter estimates may appear significantly different from zero when in fact they are not". If error terms follow the random walk, a modified model of the type in Equation 2.64 on p. 67 of the text can be used.

2.20 (a) The usual least squares estimator is $\widehat{\beta} = \frac{\Sigma x_t y_t}{\Sigma x_t^2}$, and it is unbiased, since

$E(\widehat{\beta}) = E\left(\frac{\Sigma x_t y_t}{\Sigma x_t^2} \right) = \frac{\Sigma x_t E(y_t)}{\Sigma x_t^2} = \frac{\Sigma x_t E(\beta x_t + \epsilon_t)}{\Sigma x_t^2} = \frac{\Sigma x_t \beta x_t}{\Sigma x_t^2} = \beta$. The variance of $\widehat{\beta}$ is

$$V(\widehat{\beta}) = E\left(\frac{\Sigma x_t y_t}{\Sigma x_t^2} \right) = \frac{V(\Sigma x_t y_t)}{(\Sigma x_t^2)^2} = \frac{1}{(\Sigma x_t^2)^2} \left[\Sigma x_t^2 \sigma^2 + 2 \sum_{t<u}\sum x_t x_u Cov(y_t, y_u) \right]$$

$$= \frac{\sigma^2}{(\Sigma x_t^2)} + \frac{2(n-1)\rho\sigma^2}{(\Sigma x_t^2)^2}$$

($Cov(y_t, y_u) = Corr(y_t, y_u)\sqrt{V(y_t)V(y_u)} = \rho\sigma^2$ if $t = u - 1$, and 0 otherwise, since only ϵ_t and ϵ_{t+1} are correlated).

(b) The standard deviation of $\widehat{\beta}$ in the denominator of the usual t statistic is $s \big/ \sqrt{\Sigma x_t^2}$, which will understate or overstate the true standard deviation of $\widehat{\beta}$ depending upon whether $\rho > 0$ or $\rho < 0$. This can lead to an incorrect t statistic and an incorrect test result.

2.21 (a) The usual least squares estimate is $\widehat{\beta} = \frac{\Sigma x_t y_t}{\Sigma x_t^2}$. As in Exercise 2.20(a) above, it is unbiased.

The variance of $\widehat{\beta}$ is $V(\widehat{\beta}) = E\left(\frac{\Sigma x_t y_t}{\Sigma x_t^2} \right) = \frac{V(\Sigma x_t y_t)}{(\Sigma x_t^2)^2} = \frac{\Sigma x_t^2 V(y_t)}{(\Sigma x_t^2)^2} = \frac{\Sigma x_t^2 (\sigma^2 / x_t^2)}{(\Sigma x_t^2)^2} = \frac{\sigma^2}{(\Sigma x_t^2)^2}$.

(b) $E(\widehat{\beta}^*) = E\left(\frac{\Sigma x_t^2 x_t y_t}{\Sigma x_t^2 x_t^2} \right) = \frac{\Sigma x_t^3 E(y_t)}{\Sigma x_t^4} = \frac{\Sigma x_t^3 E(\beta x_t + \epsilon_t)}{\Sigma x_t^4} = \frac{\Sigma x_t^3 \beta x_t}{\Sigma x_t^4} = \beta$ - unbiased.

(c) $V(\widehat{\beta}^*) = V\left(\frac{\Sigma x_t^2 x_t y_t}{\Sigma x_t^2 x_t^2} \right) = \frac{\Sigma x_t^6 V(y_t)}{(\Sigma x_t^4)^2} = \frac{\Sigma x_t^6 (\sigma^2 / x_t^2)}{(\Sigma x_t^4)^2} = \frac{\Sigma x_t^4 \sigma^2}{(\Sigma x_t^4)^2} = \frac{\sigma^2}{\Sigma x_t^4} \leq \frac{\sigma^2}{(\Sigma x_t^2)^2} = V(\widehat{\beta})$

(this is true since $(\Sigma x_t^2)^2 \geq \Sigma x_t^4$) .

2.21 (d) The variance stabilizing technique on p. 58-59 of the text can be applied. In the example mentioned, the transformation $g(y_t) = ln(y_t)$ would be appropriate.

2.22 $\widehat{\beta}_1 = \frac{n \Sigma x_t y_t - (\Sigma x_t)(\Sigma y_t)}{n \Sigma x_t^2 - (\Sigma x_t)^2}$ and $\widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1 \overline{x}$.

Since $\overline{y}_j$ and n_j are known for each $j = 1, ..., c$, we can find $n = \Sigma n_j$,
and $\Sigma y_t = n_1 \overline{y}_1 + n_2 \overline{y}_2 + \cdots + n_c \overline{y}_c$ and then $\overline{y} = \frac{1}{n} \Sigma y_t$.
Also, $\Sigma x_t = n_1 x_1 + n_2 x_2 + \cdots + n_c x_c$ and $\overline{x} = \frac{1}{n} \Sigma x_t$. Thus, once $\widehat{\beta}_1$ is found, $\widehat{\beta}_0$ follows.
To find $\widehat{\beta}_1$ we must still find $\Sigma x_t y_t$ and Σx_t^2 .
$n_1 \overline{y}_1$ is the sum of all y values that come from the independent x value of x_1 , and in general,
$n_j \overline{y}_j$ is the sum of all y values that come from the independent x value of x_j , so that
$\Sigma x_t y_t = x_1 n_1 \overline{y}_1 + x_2 n_2 \overline{y}_2 + \cdots + x_c n_c \overline{y}_c$.
The x_j's are all known, so that Σx_t^2 can be found. $\widehat{\beta}_1$ can now be calculated.

SOLUTIONS TO EXERCISES IN CHAPTER 3

3.1 With simple exponential smoothing, $\widehat{z}_n(1) = \widehat{z}_n(2) = \cdots = \widehat{z}_n(\ell)$ for all ℓ, so that $\widehat{z}_{30}(2) = \widehat{z}_{30}(1) = 102.5$. One form of the updating relationship for simple exponential smoothing is $\widehat{z}_{n+1}(1) = \alpha\, z_{n+1} + (1-\alpha)\,\widehat{z}_n(1)$ so that
$\widehat{z}_{31}(1) = (.1)z_{31} + (.9)\widehat{z}_{30}(1) = (.1)(105) + (.9)(102.5) = 102.75$.

3.2 (a) The prediction of total sales for the next 12 months is $\widehat{z}_{60}(1) + \widehat{z}_{60}(2) + \cdots + \widehat{z}_{60}(12) = 960$.
(n is 60 since 5 years $=$ 60 months of past data is available). As pointed out in the solution to problem 3.1, under simple exponential smoothing all $\widehat{z}_n(\ell)$ for $\ell \geq 1$ are equal. Thus, $\widehat{z}_{60}(1) = 80$, and once $z_{61} = 90$ is known, we can update to $\widehat{z}_{61}(1) = \cdots = \widehat{z}_{61}(11)$, where
$\widehat{z}_{61}(1) = (.2)(90) + (.8)(80) = 82$. The revised sales forecast for the rest of the year is $(82)(11) = 902$.

(b) The method for determining α is described in Section 3.3.2 of the text on pages 87-88.

(c) If the trend is slowly moving, simple exponential smoothing may be appropriate. Ideally, the residuals should be tested for autocorrelation to if a simple exponential smoothing model is appropriate.

3.3
$$\widehat{z}_n(2) = \alpha[z_{n+1} + (1-\alpha)z_n + (1-\alpha)^2 z_{n-1} + \cdots\ ]$$
$$= \alpha z_{n+1} + \alpha(1-\alpha)[z_n + (1-\alpha)z_{n-1} + \cdots\ ] = \alpha z_{n+1} + (1-\alpha)\widehat{z}_n(1)$$
If we substitute $\widehat{z}_n(1)$ for z_{n+1} in this expression, it becomes
$\widehat{z}_n(2) = \alpha\widehat{z}_n(1) + (1-\alpha)\widehat{z}_n(1) = \widehat{z}_n(1)$.
Suppose that $\widehat{z}_n(k) = \widehat{z}_n(1)$ and
$$\widehat{z}_n(k) = \alpha[\widehat{z}_n(k-1) + (1-\alpha)\widehat{z}_n(k-2) + (1-\alpha)^2\widehat{z}_n(k-3) + \cdots$$
$$+\ (1-\alpha)^{k-2}\widehat{z}_n(1) + (1-\alpha)^{k-1}z_n + \cdots] = \widehat{z}_n(1) \text{ for } k = 1, 2, ..., \ell - 1 .$$
Then $\widehat{z}_n(\ell) = \alpha[\widehat{z}_n(\ell-1) + (1-\alpha)\widehat{z}_n(\ell-2) + (1-\alpha)^2\widehat{z}_n(\ell-3) + \cdots$
$$+\ (1-\alpha)^{\ell-2}\widehat{z}_n(1) + (1-\alpha)^{\ell-1}z_n + \cdots]$$
$$= \alpha\widehat{z}_n(\ell-1) + (1-\alpha)\alpha[\widehat{z}_n(\ell-2) + (1-\alpha)\widehat{z}_n(\ell-3) + \cdots$$
$$+\ (1-\alpha)^{\ell-3}\widehat{z}_n(1) + (1-\alpha)^{\ell-2}z_n + \cdots]$$
$$= \alpha\widehat{z}_n(\ell-1) + (1-\alpha)\widehat{z}_n(\ell-1) = \widehat{z}_n(\ell-1) = \widehat{z}_n(1)$$

3.4 The optimal smoothing constant α (the one that minimizes SSE(α)) is $\alpha = 0$ (or $\omega = 1$), which gives the same minimum SSE(α) as the globally constant mean model.

3.5 For the smoothed statistic, the updating equation is $S_t = \alpha z_t + (1-\alpha)S_{t-1}$.
The N-observation moving average is $\overline{z}_t^{(N)} = \frac{1}{N}\,[z_t + z_{t-1} + \cdots + z_{t-N+1}]$
and at time $t-1$ it is $\overline{z}_{t-1}^{(N)} = \frac{1}{N}\,[z_{t-1} + z_{t-2} + \cdots + z_{t-N}]$.
It follows that $N\,\overline{z}_t^{(N)} = z_t + z_{t-1} + \cdots + z_{t-N+1}$ and
$N\,\overline{z}_{t-1}^{(N)} = z_{t-1} + z_{t-2} + \cdots + z_{t-N} = N\overline{z}_t^{(N)} - z_t + z_{t-N}$, or equivalently,

$\overline{z}_t^{(N)} = \overline{z}_{t-1}^{(N)} + \frac{1}{N}\,(z_t - z_{t-N})$.
$z_t = \mu + \epsilon_t$ and S_{t-1} are uncorrelated since the ϵ's are uncorrelated, and S_{t-1} depends only upon what has occurred on or before time $t-1$. Let us denote the variance of S_n by V
(assumed to be the same for all n). Then (assuming $\alpha \neq 0$),
$V[S_t] = \alpha^2\, Var[z_t] + (1-\alpha)^2\, V[S_{t-1}] \;\; \rightarrow \;\; V = \alpha^2\,\sigma^2 + (1-\alpha)^2\, V$

$$\rightarrow \;\; V = \frac{\alpha^2}{1-(1-\alpha)^2}\,\sigma^2 = \frac{\alpha}{2-\alpha}\,\sigma^2 .$$

Since $\overline{z}_t^{(N)}$ is a sample mean, its variance is $V[\overline{z}_t^{(N)}] = \frac{\sigma^2}{N}$.
In order for the variances of S_t and $\overline{z}_t^{(N)}$ to be equal, we must have $\frac{\alpha}{2-\alpha} = \frac{1}{N}$,
or equivalently, $\alpha = \frac{2}{N+1}$.

3.6 $S_n = \alpha[z_n + (1-\alpha)z_{n-1} + (1-\alpha)^2 z_{n-2} + \cdots + (1-\alpha)^{n-t_0} z_{t_0} + \cdots]$
$E(S_n) = \alpha[E(z_n) + (1-\alpha)E(z_{n-1}) + (1-\alpha)^2 E(z_{n-2}) + \cdots + (1-\alpha)^{n-t_0} E(z_{t_0}) + \cdots]$
$E(z_t) = \mu$ for all t except t_0 for which $E(z_{t_0}) = \mu + \delta$, so that
$E(S_n) = \alpha[\mu + (1-\alpha)\mu + (1-\alpha)^2\mu + \cdots + (1-\alpha)^{n-t_0}(\mu+\delta) + \cdots] = \mu + \alpha(1-\alpha)^{n-t_0}\,\delta$
(this is actually a limit assuming n is large and observations go back indefinitely into the past). If $0 < \alpha < 1$ then as $n \to \infty$, $(1-\alpha)^{n-t_0} \to 0$, so that $E(S_n) \to \mu$.

3.7 $S_n = \alpha[z_n + (1-\alpha)z_{n-1} + (1-\alpha)^2 z_{n-2} + \cdots + (1-\alpha)^{n-t_0} z_{t_0} + \cdots]$
$E(S_n) = \alpha[\mu + \delta + (1-\alpha)(\mu+\delta) + (1-\alpha)^2(\mu+\delta) + \cdots + (1-\alpha)^{n-t_0}(\mu+\delta)$
$\quad + (1-\alpha)^{n-t_0-1}\mu + (1-\alpha)^{n-t_0-2}\mu + \cdots]$
$\quad = \mu + \alpha[\delta + (1-\alpha)\delta + (1-\alpha)^2\delta + \cdots + (1-\alpha)^{n-t_0}\delta] = \mu + \delta\,[1 - (1-\alpha)^{n-t_0+1}]$
which has a limit of $\mu_1 = \mu + \delta$ as $n \to \infty$.

3.8 $S_n^{[1]} = (1-\omega) \sum_{j\geq 0} \omega^j z_{n-j} = \; = (1-\omega) \sum_{j\geq 1} \omega^{j-1} z_{n+1-j}$ (shift the index of j by 1),

$$S_n^{[2]} = (1-\omega)^2 \sum_{j\geq 0} (j+1)\omega^j z_{n-j} = (1-\omega)^2 \sum_{j\geq 1} j\omega^{j-1} z_{n+1-j} \; .$$

$$\widehat{z}_n(1) = 2S_n^{[1]} - S_n^{[2]} + \tfrac{1-\omega}{\omega}\,(S_n^{[1]} - S_n^{[2]})$$

$$= 2(1-\omega) \sum_{j\geq 1} \omega^{j-1} z_{n+1-j} - (1-\omega)^2 \sum_{j\geq 1} j\omega^{j-1} z_{n+1-j}$$

$$+ \tfrac{1-\omega}{\omega}\,[\,(1-\omega) \sum_{j\geq 1} \omega^{j-1} z_{n+1-j} - (1-\omega)^2 \sum_{j\geq 1} j\omega^{j-1} z_{n+1-j}\,]$$

$$= \sum_{j\geq 1} [\,2(1-\omega) + \tfrac{(1-\omega)^2}{\omega} - (1-\omega)^2 j - \tfrac{(1-\omega)^3}{\omega} j\,]\omega^{j-1} z_{n+1-j}$$

$$\to \quad \pi_j = [\,2(1-\omega) + \tfrac{(1-\omega)^2}{\omega} - (1-\omega)^2 j - \tfrac{(1-\omega)^3}{\omega} j\,]\omega^{j-1} \; .$$

For $\alpha = .10$ $(\omega = .90)$, $\pi_j = (.2111 - .0111j)(.9)^{j-1}$

$$\widehat{z}_n(2) = 2S_n^{[1]} - S_n^{[2]} + 2\,\tfrac{1-\omega}{\omega}\,(S_n^{[1]} - S_n^{[2]})$$

$$= 2(1-\omega) \sum_{j\geq 1} \omega^{j-1} z_{n+1-j} - (1-\omega)^2 \sum_{j\geq 1} j\omega^{j-1} z_{n+1-j}$$

$$+ 2\,\tfrac{1-\omega}{\omega}\,[\,(1-\omega) \sum_{j\geq 1} \omega^{j-1} z_{n+1-j} - 2(1-\omega)^2 \sum_{j\geq 1} j\omega^{j-1} z_{n+1-j}\,]$$

$$= \sum_{j\geq 1} [\,2(1-\omega) + \tfrac{2(1-\omega)^2}{\omega} - (1-\omega)^2 j - \tfrac{2(1-\omega)^3}{\omega} j\,]\omega^{j-1} z_{n+1-j}$$

$\rightarrow\ \pi_j^{(2)} = [\,2(1-\omega) + \frac{2(1-\omega)^2}{\omega} - (1-\omega)^2 j - \frac{2(1-\omega)^3}{\omega} j\,]\omega^{j-1}$
For $\alpha = .10$ ($\omega = .90$), $\pi_j = (.2222 - .0122j)(.9)^{j-1}$.

3.9 (a) $\omega_1 = 1 - \alpha_1 = .8$ and $\omega_2 = 1 - \alpha_2 = .9$. Using Equation 3.41 of the text with $n = 24$, $\widehat{\mu}_n = \widehat{\beta}_{0,n} = 30$ and $\widehat{\beta}_n = \widehat{\beta}_{1,n} = 2$ we get
$\widehat{\beta}_{0,n+1} = \widehat{\mu}_{n+1} = \widehat{\mu}_{25} = (1-\omega_1)z_{n+1} + \omega_1(\widehat{\mu}_n + \widehat{\beta}_n) = .2(28) + (.8)(30+2) = 31.2$ and
$\widehat{\beta}_{1,n+1} = \widehat{\beta}_{n+1} = \widehat{\beta}_{25} = (1-\omega_2)(\widehat{\mu}_{n+1} - \widehat{\mu}_n) + \omega_2\widehat{\beta}_n = (.1)(31.2-30) + (.9)(2) = 1.92$.
The original forecasts were (using Equation 3.38) $\widehat{z}_n(1) = \widehat{\beta}_{0,n} + \widehat{\beta}_{1,n} = 32$ (the forecast of z_{25}), $\widehat{z}_n(2) = \widehat{\beta}_{0,n} + 2\widehat{\beta}_{1,n} = 34$ (the forecast of z_{26}) and $\widehat{z}_n(3) = \widehat{\beta}_{0,n} + 3\widehat{\beta}_{1,n} = 36$ (the forecast of z_{27}), and the updated forecasts are (using Equation 3.42)
$\widehat{z}_{n+1}(1) = \widehat{\mu}_{n+1} + \widehat{\beta}_{n+1} = 31.2 + 1.92 = 33.12$ (the forecast of z_{26}) and
$\widehat{z}_{n+1}(2) = \widehat{\mu}_{n+1} + 2\widehat{\beta}_{n+1} = 31.2 + 2(1.92) = 35.04$ (the forecast of z_{27}).
Since z_{n+2} ($= z_{26}$) and z_{n+3} ($= z_{27}$) are known, it is possible to repeat the procedure to get $\widehat{\mu}_{n+2}$ and $\widehat{\beta}_{n+2}$ and updated forecasts $\widehat{z}_{n+2}(1)$, etc.

3.9 (b) Using Equation 3.40 with $n = 24$, we get updated values
$\widehat{\beta}_{0,25} = \alpha^2 z_{25} + (1-\omega^2)(\widehat{\beta}_{0,24} + \widehat{\beta}_{1,24}) = (.19)(28) + (.9)^2(30+2) = 31.24$ and

$\widehat{\beta}_{1,25} = \frac{1-\omega}{1+\omega}(\widehat{\beta}_{0,25} - \widehat{\beta}_{0,24}) + \frac{2\omega}{1+\omega}\widehat{\beta}_{1,24} = \frac{.1}{1.9}(31.24-30) + \frac{2(.9)}{1.9}(2) = 1.96$.
The updated forecasts are then found using $\widehat{z}_{25}(\ell) = \widehat{\beta}_{0,25} + \widehat{\beta}_{1,25}\,\ell$, so that
$\widehat{z}_{25}(1) = \widehat{\beta}_{0,25} + \widehat{\beta}_{1,25} = 31.24 + 1.96 = 33.20$ (the forecast of z_{26}) and
$\widehat{z}_{25}(2) = \widehat{\beta}_{0,25} + 2\widehat{\beta}_{1,25} = 31.24 + 2(1.96) = 35.16$.

From Equation 3.37 we have $\widehat{\beta}_{0,n} = \widehat{\beta}_{0,24} = 2S_{24}^{[1]} - S_{24}^{[2]} \ \rightarrow\ 2S_{24}^{[1]} - S_{24}^{[2]} = 30$, and

$\widehat{\beta}_{1,n} = \widehat{\beta}_{1,24} = \frac{1-\omega}{\omega}(S_{24}^{[1]} - S_{24}^{[2]}) \ \rightarrow\ \frac{.1}{.9}(S_{24}^{[1]} - S_{24}^{[2]}) = 2$. Solving these two equations for

$S_{24}^{[1]}$ and $S_{24}^{[2]}$ results in $S_{24}^{[1]} = 12$ and $S_{24}^{[2]} = -6$.
Then from Equation 3.34, $S_{25}^{[1]} = (1-\omega)z_{25} + \omega S_{24}^{[1]} = (.1)(28) + (.9)(12) = 13.6$
$S_{25}^{[2]} = (1-\omega)S_{25}^{[1]} + \omega S_{24}^{[2]} = (.1)(13.6) + (.9)(-6) = -4.04$,
$S_{26}^{[1]} = (1-\omega)z_{26} + \omega S_{25}^{[1]} = (.1)(31) + (.9)(13.6) = 15.34$
$S_{26}^{[2]} = (1-\omega)S_{26}^{[1]} + \omega S_{25}^{[2]} = (.1)(15.34) + (.9)(-4.04) = -2.102$, and
$S_{27}^{[1]} = (1-\omega)z_{27} + \omega S_{26}^{[1]} = (.1)(36) + (.9)(15.34) = 17.406$
$S_{27}^{[2]} = (1-\omega)S_{27}^{[1]} + \omega S_{26}^{[2]} = (.1)(17.406) + (.9)(-2.102) = -.1512$.

3.10 We are given $\omega = .9$ ($\alpha = .1$) and σ is known . Then from Equation 3.62, $V[e_n(\ell)] = \sigma^2 c_\ell^2$,
where $c_\ell^2 = 1 + \frac{1-\omega}{(1+\omega)^3}[\,(1+4\omega+5\omega^2) + 2\ell(1-\omega)(1+3\omega) + 2\ell^2(1-\omega)^2\,]$, so that

$V[e_n(1)] = \sigma^2 c_1^2 = \frac{.1}{(1.9)^2}[\,(1 + 4(.9) + 5(.9)^2 + 2(.1)(3.7) + 2(.1)^2]\,\sigma^2 = .2607\sigma^2$,

$V[e_n(2)] = .2828\sigma^2$, $V[e_n(3)] = .3061\sigma^2$ and $V[e_n(4)] = .3305\sigma^2$.
The 95% prediction intervals are:
for 1-step-ahead $\widehat{z}_n(1) \pm 1.96\,\sigma\sqrt{.2607}$ which becomes $\widehat{z}_n(1) \pm 1.00\sigma$,
for 2-step-ahead $\widehat{z}_n(2) \pm 1.042\sigma$, for 3-step-ahead $\widehat{z}_n(3) \pm 1.084\sigma$, and
for 4-step-ahead $\widehat{z}_n(1) \pm 1.127\sigma$.
It is assumed that the mean absolute deviation $\widehat{\Delta}_e$ is known. The estimate for σ_e is then $\widehat{\sigma}_e \cong 1.25\,\widehat{\Delta}_e = \widehat{\sigma}\sqrt{c_1}$.

Equation 3.68 on page 130 of the text gives the prediction interval for a sum of K future observations. For the case of double exponential smoothing, the matrix $\mathbf{F}^{-1}$ is

$$\mathbf{F}^{-1} = \begin{bmatrix} 1-\omega^2 & (1-\omega)^2 \\ (1-\omega)^2 & \frac{(1-\omega)^3}{\omega} \end{bmatrix} \quad \text{and} \quad \mathbf{F}_* = \begin{bmatrix} \frac{1}{1-\omega^2} & \frac{-\omega^2}{(1-\omega^2)^2} \\ \frac{-\omega^2}{(1-\omega^2)^2} & \frac{\omega^2(1+\omega^2)}{(1-\omega^2)^3} \end{bmatrix}$$

(page 129 of the text).

3.10 (cont'd) With $\alpha = .1$ ($\omega = .9$) these matrices become

$$\mathbf{F}^{-1} = \begin{bmatrix} .19 & .01 \\ .01 & .00111 \end{bmatrix} \quad \text{and} \quad \mathbf{F}_* = \begin{bmatrix} 5.2632 & -22.4377 \\ -22.4377 & 213.748 \end{bmatrix} \quad \text{and}$$

$$\mathbf{F}^{-1}\mathbf{F}_*\mathbf{F}^{-1} = \begin{bmatrix} .1261 & .005394 \\ .005394 & .0002916 \end{bmatrix}$$

$$\mathbf{f}(\ell) = \begin{bmatrix} 1 \\ \ell \end{bmatrix} \;\rightarrow\; \sum_{\ell=1}^{K}\mathbf{f}(\ell) = \sum_{\ell=1}^{K}\begin{bmatrix} 1 \\ \ell \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \cdots + \begin{bmatrix} 1 \\ K \end{bmatrix} = \begin{bmatrix} K \\ \frac{K(K+1)}{2} \end{bmatrix}$$

For $K = 4$, $\left(\sum_{\ell=1}^{4}\mathbf{f}'(\ell)\right)\mathbf{F}^{-1}\mathbf{F}_*\mathbf{F}^{-1}\left(\sum_{\ell=1}^{4}\mathbf{f}(\ell)\right) = \mathbf{f}'(1)\,\mathbf{F}^{-1}\mathbf{F}_*\mathbf{F}^{-1}\,\mathbf{f}(1)$

$$= [\,4 \quad 10\,]\begin{bmatrix} .1261 & .005394 \\ .005394 & .0002916 \end{bmatrix}\begin{bmatrix} 4 \\ 10 \end{bmatrix} = 2.478\,,$$

The 95% prediction interval for the sum of the next four observations is then

$\sum_{\ell=1}^{4}\widehat{z}_n(\ell) \pm 1.96\frac{\widehat{\sigma}_e}{c_1}\,[\,4 + 2.478\,]^{1/2}$. The same comment that $\widehat{\sigma}_e \cong 1.25\widehat{\Delta}_e$ applies.

The average of the next four observations is $\frac{1}{4}\sum_{\ell=1}^{4}\widehat{z}_n(\ell)$, and has a standard deviation

which is $\frac{1}{4}$ of that of $\sum_{\ell=1}^{4}\widehat{z}_n(\ell)$, so that the prediction interval for the average is

$$\frac{1}{4}\sum_{\ell=1}^{4}\widehat{z}_n(\ell) \pm 1.96\frac{\widehat{\sigma}_e}{c_1}\cdot\frac{1}{4}\,[\,4 + 2.478\,]^{1/2}$$

3.11 $\sum_{j\geq 0} j\omega^j = \omega\frac{\partial}{\partial\omega}\sum_{j\geq 0}\omega^j = \omega\frac{\partial}{\partial\omega}\left[\,\frac{1}{1-\omega}\,\right] = \frac{\omega}{(1-\omega)^2}$,

$$\sum_{j\geq 0} j^2\omega^j = \omega\frac{\partial}{\partial\omega}\sum_{j\geq 0} j\omega^j = \omega\frac{\partial}{\partial\omega}\frac{\omega}{(1-\omega)^2} = \frac{\omega[(1-\omega)^2+2\omega(1-\omega)]}{(1-\omega)^3} = \frac{\omega(1+\omega)}{(1-\omega)^3},$$

$$\sum_{j\geq 0} j^3\omega^j = \omega\frac{\partial}{\partial\omega}\sum_{j\geq 0} j^2\omega^j = \omega\frac{\partial}{\partial\omega}\frac{\omega(1+\omega)}{(1-\omega)^3}$$

$$= \frac{\omega[(1-\omega)^3(1+2\omega)+3\omega(1+\omega)(1-\omega)^2]}{(1-\omega)^6} = \frac{\omega(1+4\omega+\omega^2)}{(1-\omega)^4}$$

$$\sum_{j\geq 0} j^4\omega^j = \omega\frac{\partial}{\partial\omega}\sum_{j\geq 0} j^3\omega^j = \omega\frac{\partial}{\partial\omega}\frac{\omega(1+4\omega+\omega^2)}{(1-\omega)^4}$$

$$= \frac{\omega[(1-\omega)^4(1+8\omega+3\omega^2)+4\omega(1+4\omega+\omega^2)(1-\omega)^3]}{(1-\omega)^8} = \frac{\omega(1+11\omega+11\omega^2+\omega^3)}{(1-\omega)^5}$$

3.12 The smoothed statistics are calculated from on the relationship $S_n^{[k]} = (1-\omega)S_n^{[k-1]} + \omega S_{n-1}^{[k]}$. We use proof by induction on k. Suppose that the statement $S_n^{[r]} = \frac{(1-\omega)^r}{(r-1)!}\sum_{j\geq 0}\left[\prod_{i=1}^{r-1}(j+i)\right]\omega^j z_{n-j}$ is true for $r = 1, 2, ..., k$ and for all n.

We wish to show that the statement is true for $r = k+1$.

But $S_n^{[k+1]} = (1-\omega)S_n^{[k]} + \omega S_{n-1}^{[k+1]} = (1-\omega)S_n^{[k]} + \omega(1-\omega)S_{n-1}^{[k]} + \omega^2 S_{n-2}^{[k+1]}$

$$= (1-\omega)S_n^{[k]} + \omega(1-\omega)S_{n-1}^{[k]} + \omega^2(1-\omega)S_{n-2}^{[k]} + \cdots$$

$$= (1-\omega)\frac{(1-\omega)^k}{(k-1)!}\sum_{j\geq 0}\left[\prod_{i=1}^{k-1}(j+i)\right]\omega^j z_{n-j}$$

$$+ \omega(1-\omega)\frac{(1-\omega)^k}{(k-1)!}\sum_{j\geq 0}\left[\prod_{i=1}^{k-1}(j+i)\right]\omega^j z_{n-1-j}$$

$$+ \omega^2(1-\omega)\frac{(1-\omega)^k}{(k-1)!}\sum_{j\geq 0}\left[\prod_{i=1}^{k-1}(j+i)\right]\omega^j z_{n-2-j} + \cdots$$

$$= \frac{(1-\omega)^{k+1}}{(k-1)!}\sum_{j\geq 0}\left[\prod_{i=1}^{k-1}(j+i)\right]\omega^j z_{n-j}$$

$$+ \frac{(1-\omega)^{k+1}}{(k-1)!}\sum_{j\geq 0}\left[\prod_{i=1}^{k-1}(j+i)\right]\omega^{j+1} z_{n-1-j}$$

$$+ \frac{(1-\omega)^{k+1}}{(k-1)!}\sum_{j\geq 0}\left[\prod_{i=1}^{k-1}(j+i)\right]\omega^{j+2} z_{n-2-j} + \cdots$$

$$= \frac{(1-\omega)^{k+1}}{(k-1)!}\sum_{j\geq 0}\left[\prod_{i=1}^{k-1}(j+i)\right]\omega^j z_{n-j}$$

$$+\ \tfrac{(1-\omega)^{k+1}}{(k-1)!} \sum_{j\geq 1} \left[\prod_{i=1}^{k-1} (j-1+i) \right] \omega^j z_{n-j}$$

$$+\ \tfrac{(1-\omega)^{k+1}}{(k-1)!} \sum_{j\geq 2} \left[\prod_{i=1}^{k-1} (j-2+i) \right] \omega^j z_{n-j} \quad + \cdots$$

$$= \sum_{r\geq 0} \tfrac{(1-\omega)^{k+1}}{(k-1)!} \sum_{j\geq r} \left[\prod_{i=1}^{k-1} (j-r+i) \right] \omega^j z_{n-j}$$

$$= \quad \tfrac{(1-\omega)^{k+1}}{(k-1)!} \sum_{j\geq 0} \left[\sum_{r=0}^{j} \left[\prod_{i=1}^{k-1} (j-r+i) \right] \right] \omega^j z_{n-j}$$

3.12 (cont'd) If we can show that $\sum_{r=0}^{j} \left[\prod_{i=1}^{k-1} (j-r+i) \right] = \frac{1}{k} \prod_{i=1}^{k} (j+i)$ (Equation A)

then the result follows from mathematical induction.
Doing the change of variable $t = j - r$, Equation A becomes

$\sum_{t=0}^{j} \left[\prod_{i=1}^{k-1} (t+i) \right] = \frac{1}{k} \prod_{i=1}^{k} (j+i)$. This equation can be established by induction on j.

It is true for $j = 0$ and for any value of k, since for $j = 0$, the left-hand-side is $(k-1)!$, and the right-hand-side is $\frac{1}{k}\, k! = (k-1)!$. Suppose that the statement is true for all values of j from $j = 0, 1, 2, ..., s$. We wish to show that it is true for $s + 1$.

$$\sum_{t=0}^{s+1} \left[\prod_{i=1}^{k-1} (t+i) \right] = \sum_{t=0}^{s} \left[\prod_{i=1}^{k-1} (t+i) \right] + \prod_{i=1}^{k-1} (s+1+i)$$

$$= \frac{1}{k} \prod_{i=1}^{k} (s+i) + \prod_{i=1}^{k-1} (s+1+i) = \frac{1}{k} \prod_{i=0}^{k-1} (s+1+i) + \prod_{i=1}^{k-1} (s+1+i)$$

$$= \prod_{i=1}^{k-1} (s+1+i) \left[\tfrac{s+1}{k} + 1 \right] = \frac{1}{k} \prod_{i=1}^{k} (s+1+i)$$, which is the statement for $s + 1$.

This establishes the validity of Equation A and completes the result.

3.13 (a) True. As pointed out on page 86 of the text, under simple exponential smoothing, the forecasts from a fixed forecast origin n are the same for all lead times ℓ.

(b) True. From Equation 3.38 on page 106 of the text, we see that $\widehat{z}_n(\ell)$ is a linear function of ℓ.

(c) True. From Equation 3.60 on page 128 of the text we see that the prediction intervals under simple exponential smoothing are the same for all ℓ.

SOLUTIONS TO EXERCISES IN CHAPTER 4

4.1 (a) In this case $k = 0$ for the trend component and $m = 1$ for the seasonal component. The transition matrix $\mathbf{L}$ is a 3 x 3 matrix of the form (see pages 153-154 of the text) $\begin{bmatrix} \mathbf{L_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{L_2} \end{bmatrix}$ where $\mathbf{L_1}$ is the 1 x 1 transition matrix for the trend component, [1] ,
and $\mathbf{L_2}$ is the 2 x 2 matrix

$$\mathbf{L_2} = \begin{bmatrix} \cos\frac{\pi}{2} & \sin\frac{\pi}{2} \\ -\sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \text{ so that } \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

(b) $k = 1$ for the trend component and $m = 1$ for the seasonal component. The transition matrix $\mathbf{L}$ is the 4 x 4 matrix $\begin{bmatrix} \mathbf{L_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{L_2} \end{bmatrix}$, where $\mathbf{L_1}$ is the 2 x 2 transition matrix for the trend component (see p. 97 of the text), $\mathbf{L_1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{L_2}$ is the 2 x 2 matrix as in part (a),

$$\mathbf{L_2} = \begin{bmatrix} \cos\frac{\pi}{2} & \sin\frac{\pi}{2} \\ -\sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \text{ so that } \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

(b) $k = 1$ for the trend component and $m = 2$ for the seasonal component. The transition matrix $\mathbf{L}$ is the 4 x 4 matrix $\begin{bmatrix} \mathbf{L_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{L_2} \end{bmatrix}$, where $\mathbf{L_1}$ is the 2 x 2 transition matrix as in part (b),

$$\mathbf{L_1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \text{ and } \mathbf{L_2} \text{ is the 4 x 4 matrix } \mathbf{L_2} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \text{ so that}$$

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}.$$

SOLUTIONS TO EXERCISES IN CHAPTER 5

5.1 (a) The slowly decaying sample autocorrelation function of the z_t's and the more rapidly exponential decay in the SACF of the differenced series indicates that the z_t series is non-stationary but the series of first differences is stationary. A possible choice for the model of the differenced series is AR(1) (MA would require the autocorrelations to be very near 0 after a few lags).

(b) The SACF seems to be a damped sine wave, which indicates an AR(2) model (or ARMA(2,1) or ARMA(2,2) model).

5.2 A stationary model requires (among other things) $E(z_t) = \mu$ for all t. This is not the case for this model. This model is not stationary in the mean.

5.3 (a) The z_t's are not stationary, but the series of second differences is stationary with an MA(1) model.

(b) The model is invertible since the sum of the $|\pi_k|$'s is finite (they form an infinite geometric progression with $|\theta| < 1$ - see part (d) following).

(c) The second differences form a MA(1) model so that for te second difference series, ρ_1 is $\frac{-\theta}{1+\theta^2}$ and for $k \geq 2$, $\rho_k = 0$.

(d) The ψ weights are found from $(1 + \psi_1 B + \psi_2 B^2 + \cdots)(1 - B)^2 = 1 - \theta B$, or equivalently
$(1 + \psi_1 B + \psi_2 B^2 + \cdots)(1 - 2B + B^2) = 1 - \theta B$, so that
B: $\psi_1 - 2 = -\theta \quad \rightarrow \quad \psi_1 = 2 - \theta$
B^2: $\psi_2 - 2\psi_1 + 1 = 0 \quad \rightarrow \quad \psi_2 = 2\psi_1 - 1 = 2(2-\theta) - 1 = 3 - 2\theta$
B^3: $\psi_3 - 2\psi_2 + \psi_1 = 0 \quad \rightarrow \quad \psi_3 = 2\psi_2 - \psi_1 = 4 - 3\theta$
B^k: $\psi_k - 2\psi_{k-1} + \psi_{k-2} = 0 \quad \rightarrow \quad \psi_k = 2\psi_{k-1} - \psi_{k-2} = k + 1 - k\theta$

The π weights are found from $(1 - B)^2 = (1 - \theta B)(1 + \pi_1 B + \pi_2 B^2 + \cdots)$, or equivalently,
$1 - 2B + B^2 = (1 - \theta B)(1 + \pi_1 B + \pi_2 B^2 + \cdots)$, so that
B: $\pi_1 - \theta = -2 \quad \rightarrow \quad \pi_1 = \theta - 2$
B^2: $\pi_2 - \theta\pi_1 = 1 \quad \rightarrow \quad \pi_2 = \theta\pi_1 + 1 = \theta(\theta - 2) + 1$
B^3: $\pi_3 - \theta\pi_2 = 0 \quad \rightarrow \quad \pi_3 = \theta\pi_2 = \theta^2(\theta - 2) + \theta$
B^k: $\pi_k = \theta\pi_{k-1} \quad \rightarrow \quad \pi_k = \theta\pi_{k-1} = \theta^{k-1}(\theta - 2) + \theta^{k-2}$.

(e) We can use Equation 5.62 on p. 239 to get $z_t(1) = \psi_1 a_t + \psi_2 a_{t-1} + \cdots$, where the ψ's were found in part (d) above.
Alternatively, we can write $z_{t+1} = 2z_t - z_{t-1} + a_{t+1} - \theta a_t$ so that
$z_t(1) = E(z_{t+1} \mid z_t, z_{t-1}, ...) = 2z_t - z_{t-1} - \theta a_t$ (since $E(a_{t+1} \mid z_t, z_{t-1}, ...) = 0$).

$e_t(1) = a_{t+1}$, which has variance σ^2 (this is true for the 1-step-ahead forecast for any model).

5.4 (a) $y_t = x_t + z_t = \phi x_{t-1} + a_t + z_t$, but $y_{t-1} = x_{t-1} + z_{t-1}$, so that
$x_{t-1} = y_{t-1} - z_{t-1}$ and $y_t = \phi(y_{t-1} - z_{t-1}) + a_t + z_t = \phi y_{t-1} + a_t + z_t - \phi z_{t-1}$
Alternatively, the model $(1-\phi B)x_t = a_t$ becomes $(1-\phi B)(y_t - z_t) = a_t$, or equivalently,
$(1-\phi B)y_t = a_t + (1-\phi B)z_t \;\; \rightarrow \;\; y_t = \phi y_{t-1} + a_t + z_t - \phi z_{t-1}$.
Then, $y_t = \phi[\phi y_{t-2} + a_{t-1} + z_{t-1} - \phi z_{t-2}] + a_t + z_t - \phi z_{t-1}$
$= \phi^2 y_{t-2} + a_t + \phi a_{t-1} + z_t - \phi^2 z_{t-2}$.
Continuing in this way, we get
$y_t = \phi^k y_{t-k} + a_t + \phi a_{t-1} + \cdots + \phi^{k-1} a_{t-k+1} + z_t - \phi_k z_{t-k}$, so that if $|\phi| < 1$, then as $k \rightarrow \infty$ we have $y_t = a_t + \phi a_{t-1} + \cdots + z_t$. Since the z's are uncorrelated with the a's, the variance of y_t is the same as that for the AR(1) model plus $V(z_t) = \sigma_z^2$,

i.e., $V(y_t) = \frac{\sigma_a^2}{1-\phi^2} + \sigma_z^2$. Also, the covariance k-steps-apart for $k \geq 1$ is the same as for the

AR(1) model because the z's and a's are uncorrelated, i.e., Cov$(y_{t+k}, y_t) = \frac{\phi^k \sigma_a^2}{1-\phi^2}$, so that

$$\rho_k = \left(\frac{\phi^k \sigma_a^2}{1-\phi^2}\right) \Big/ \left(\frac{\sigma_a^2}{1-\phi^2} + \sigma_z^2\right) = \frac{\phi^k}{1+(1-\phi^2)\frac{\sigma_z^2}{\sigma_a^2}} = \phi^{k-1}\rho_1 \; .$$

(b) The ARMA(1,1) model has autocorrelation function $\rho_k = \phi^{k-1}\rho_1$ which is the same form as the model in part (a).

5.5 (a) If the model is represented as $z_t = \phi_{k1} z_{t-1} + \phi_{k2} z_{t-2} + \cdots + \phi_{kk} z_{t-k} + a_t$, then the partial autocorrelation for lag k for is ϕ_{kk} .

(b) (i) This is an AR(2) model with $\phi_1 = 1.2$ and $\phi_2 = -.8$. Using the representations for the partial autocorrelations as given on p. 210 of the text we have $\phi_{11} = \rho_1 = \frac{\phi_1}{1-\phi_2} = \frac{2}{3}$.

Also, $\rho_2 = \frac{\phi_1^2}{1-\phi_2} + \phi_2 = 0$, so that $\phi_{22} = \frac{\rho_2 - \rho_1^2}{1-\rho_1^2} = -.8$.

For the AR(2) model, $\phi_{kk} = 0$ for $k \geq 3$.

(ii) This is an AR(1) model with $\phi = .7$. Then $\phi_{11} = \rho_1 = \phi = .7$ and $\phi_{kk} = 0$ for $k \geq 2$.

(iii) This is an ARMA(1,1) model with $\phi = .7$ and $\theta = .5$. From p. 222 of the text we have that $\phi_{11} = \rho_1 = \frac{(1-\phi\theta)(\phi-\theta)}{1+\theta^2-2\phi\theta} = .2364$, and ϕ_{kk} for $k \geq 1$ is as for the MA(1) model (as

found n page 218 of the text) so that $\phi_{22} = \frac{-\theta^2(1-\theta^2)}{1-\theta^6} = -.1905$ and

$\phi_{33} = \frac{-\theta^3(1-\theta^2)}{1-\theta^8} = -.0941$.

(iv) This is an MA(1) model with $\theta = .5$. The PACF for the MA(1) is found on page 218 of

the text. $\phi_{11} = \frac{-\theta(1-\theta^2)}{1-\theta^4} = -.4$, and $\phi_{22} = -.1905$, $\phi_{33} = -.0941$ are the same as in part (iii) above.

5.6 $\sum_{t=1}^{n} z_t = z_1 + z_2 + \cdots + z_n = a_1 - a_0 + a_2 - a_1 + \cdots + a_n - a_{n-1} = a_n - a_0$ so that the sample mean is $\frac{1}{n}\sum_{t=1}^{n} z_t = \frac{a_n - a_0}{n}$ and

$$V\left(\frac{1}{n}\sum_{t=1}^{n} z_t\right) = V\left(\frac{a_n + a_0}{n}\right) = \frac{1}{n^2}\,V(a_n + a_0) = \frac{1}{n^2}(\sigma^2 + \sigma^2) = \frac{2\sigma^2}{n^2}\ .$$

Since the form of the model has a's canceling out when successive z's are added, the total random component remains small in the sum.

5.7 $z_t - z_{t-1} = a_t$

$\rightarrow\ {}_t\triangle_m = z_t - z_{t-m} = (z_t - z_{t-1}) + (z_{t-1} - z_{t-2}) + \cdots (z_{t-m+1} - z_{t-m})$
$= a_t + a_{t-1} + \cdots + a_{t-m+1} \quad\rightarrow\quad V({}_t\triangle_m) = \sigma^2 + \sigma^2 + \cdots + \sigma^2 = m\sigma^2$

${}_t\triangle^*_m = \frac{1}{m}\,[\,(z_t - z_{t-m}) + (z_{t+1} - z_{t-m+1}) + \cdots + (z_{t+m-1} - z_{t-1})]$
$= \frac{1}{m}\,[(a_t + a_{t-1} + \cdots + a_{t-m+1}) + (a_{t+1} + a_t + \cdots + a_{t-m+2})$
$+ (a_{t+m-1} + a_{t+m-2} + \cdots + a_t)]$
$= \frac{1}{m}\,[a_{t-m+1} + 2a_{t-m+2} + 3a_{t-m+3} + \cdots + (m-1)a_{t-1}$
$+ ma_t + (m-1)a_{t+1} + \cdots + 2a_{t+m-2} + a_{t+m-1}]\ .$

Then, $V({}_t\triangle^*_m) = \frac{1}{m^2}\,[\sigma^2 + 2^2\sigma^2 + 3^2\sigma^2 + \cdots + (m-1)^2\sigma^2$
$+ m^2\sigma^2 + (m-1)^2\sigma^2 + \cdots + 2^2\sigma^2 + \sigma^2]$

$\rightarrow\quad V({}_t\triangle^*_m) = \frac{\sigma^2}{m^2}\,[2(1) + 2(2^2) + 2(3^2) + \cdots + 2(m-1)^2 + m^2]$

$$= \frac{\sigma^2}{m^2}\,[\,2\left(\sum_{i=1}^{m} i^2\right) - m^2] = \frac{\sigma^2}{m^2}\,[\,2\frac{m(m+1)(2m+1)}{6} - m^2] = \frac{\sigma^2}{m^2}\,[\,\frac{2m^3+m}{3}\,]\ .$$

Then $V({}_t\triangle^*_m)\Big/V({}_t\triangle_m) = \frac{\sigma^2}{m^2}\,[\,\frac{2m^3+m}{3}\,]\Big/m\sigma^2 = \frac{2}{3} + \frac{1}{m^2}$, which has a limit of $\frac{2}{3}$ as $m \rightarrow \infty$

5.9 The Portmanteau test for model adequacy gives a value of Q^* of
$Q^* = 99[(-.32)^2 + (.20)^2 + (.05)^2 + (-.06)^2 + (-.08)^2] = 15.34$.
The chi-square distribution has $K - p - q = 5 - 1 = 4$ (assuming a mean wasn't estimated).
At the 5% level, we have $\chi^2_{.05}(4) = 9.49$. The large value of Q^* indicates that the model is not adequate.

5.10 (a) (1) AR(1) model which has forecast equation (Equation 5.66 on page 241 of the text)
$z_n(\ell) = \mu + \phi[z_n(\ell-1) - \mu] = \mu + \phi^\ell(z_n - \mu)$.
(2) AR(2) model which has forecast equation (Equation 5.72, p.
243) $z_n(\ell) = \phi_1 z_n(\ell-1) + \phi_2 z_n(\ell-2)$ (assuming $\mu = 0$) .
(3) ARIMA(1,1,1) model which has (Equation 5.78 on p.245)
$z_n(\ell) = \theta_0 + (1+\phi)z_n(\ell-1) - \phi z_n(\ell-2)$.
(4) ARIMA(0,2,3) model with $z_t = 2z_{t-1} - z_{t-2} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3}$
so that $z_n(1) = 2z_n - z_{n-1} - \theta_1 a_n - \theta_2 a_{n-1} - \theta_3 a_{n-2}$
$z_n(2) = 2z_n(1) - z_n - \theta_2 a_n - \theta_3 a_{n-1}$
$z_n(3) = 2z_n(2) - z_n(1) - \theta_3 a_n$ and $z_n(\ell) = 2z_n(\ell-1) - z_n(\ell-2)$ for $\ell \geq 4$.

5.10 (b) (1) The derivations for $\ell = 1$ and 2 for the AR(1) model are given on page 241 of the text.
$e_n(3) = z_{n+3} - z_n(3) = \mu + \phi(z_{n+2} - \mu) + a_{n+3} - [\mu + \phi^3(z_n - \mu)]$
$= \phi(z_{n+2} - \mu) + a_{n+3} - \phi^3(z_n - \mu) = \phi[\phi(z_{n+1} - \mu) + a_{n+2}] + a_{n+3} - \phi^3(z_n - \mu)$
$= \phi^2(z_{n+1} - \mu) + a_{n+3} + \phi a_{n+2} - \phi^3(z_n - \mu)$
$= \phi^2[\phi(z_n - \mu) + a_{n+1}] + a_{n+3} + \phi a_{n+2} - \phi^3(z_n - \mu) = a_{n+3} + \phi a_{n+2} + \phi^2 a_{n+1}$

Then, $V[e_n(3)] = \sigma^2[1 + \phi^2 + \phi^4] = \sigma^2\,\frac{1-\phi^6}{1-\phi^2}$ (as in Equation 5.68 on page 241).

(2) For any model, $V[e_n(\ell)] = \sigma^2(1 + \psi_1^2 + \psi_2^2 + \cdots + \psi_{\ell-1}^2)$. For the AR(2) model, we have $\psi_1 = \phi_1$ and $\psi_2 = \phi_1^2 + \phi_2$. Thus, $V[e_n(1)] = \sigma^2$, $V[e_n(2)] = \sigma^2(1 + \phi_1^2)$ and $V[e_n(3)] = \sigma^2[\,1 + \phi_1^2 + (\phi_1^2 + \phi_2)^2]$.

(3) For any model, $e_n(1) = a_{n+1}$ so that $V[e_n(1)] = \sigma^2$. For simplicity we assume $\theta_0 = 0$.
For the ARIMA(1,1,1) model, $z_{n+2} = (1+\phi)z_{n+1} - \phi z_n + a_{n+2} - \theta a_{n+1}$ so that
$e_n(2) = z_{n+2} - z_n(2) = (1+\phi)z_{n+1} - \phi z_n + a_{n+2} - \theta a_{n+1} - [(1+\phi)z_n(1) - \phi z_n]$
(this uses the formulation for $z_n(2)$ for the ARIMA(1,1,1) model on page 245) , then using the formulation for z_{n+1} and $z_n(1)$ on page 245, we have

$$\begin{aligned} e_n(2) &= (1+\phi)[(1+\phi)z_n - \phi z_{n-1} + a_{n+1} - \theta a_n] + a_{n+2} - \theta a_{n+1} \\ &\qquad - (1+\phi)[(1+\phi)z_n - \phi z_{n-1} - \theta a_n] \\ &= a_{n+2} + (1+\theta-\phi)a_{n+1} \text{ and } V[e_n(2)] = \sigma^2\,[1 + (1+\theta-\phi)^2] . \end{aligned}$$

For $\ell = 3$ we have $e_n(3) = z_{n+3} - z_n(3)$.
We first write $z_{n+3} = (1+\phi)z_{n+2} - \phi z_{n+1} + a_{n+3} - \theta a_{n+2}$, and then using the model form for z_{n+2} and z_{n+1}, we have $z_{n+2} = (1+\phi)z_{n+1} - \phi z_n + a_{n+2} - \theta a_{n+1}$
and $z_{n+1} = (1+\phi)z_n - \phi z_{n-1} + a_{n+1} - \theta a_n$ so that

$$\begin{aligned} z_{n+3} &= (1+\phi)[(1+\phi)z_{n+1} - \phi z_n + a_{n+2} - \theta a_{n+1}] \\ &\qquad - \phi[(1+\phi)z_n - \phi z_{n-1} + a_{n+1} - \theta a_n] + a_{n+3} - \theta a_{n+2} \\ &= (1+\phi)^2 z_{n+1} - 2(1+\phi)\phi z_n + \phi^2 z_{n-1} \\ &\qquad + a_{n+3} + (1+\phi-\theta)a_{n+2} - [(1+\phi)\theta + \phi]a_{n+1} + \phi\theta a_n \\ &= (1+\phi)^2[(1+\phi)z_n - \phi z_{n-1} + a_{n+1} - \theta a_n] - 2(1+\phi)\phi z_n \\ &\qquad + \phi^2 z_{n-1} + a_{n+3} + (1+\phi-\theta)a_{n+2} - [(1+\phi)\theta + \phi]a_{n+1} + \phi\theta a_n \\ &= [(1+\phi)^3 - 2\phi(1+\phi)]z_n + [\phi^2 - (1+\phi)^2\phi^2]z_{n-1} + a_{n+3} + (1+\phi-\theta)a_{n+2} \\ &\qquad + [(1+\phi)^2 - (1+\phi)\theta - \phi]a_{n+1} + \phi\theta a_n \end{aligned}$$

In a similar way,

$$\begin{aligned} z_n(3) &= (1+\phi)z_n(2) - \phi z_n(1) \\ &= (1+\phi)[(1+\phi)z_n(1) - \phi z_n] - \phi[(1+\phi)z_n - \phi z_{n-1} - \theta a_n] \\ &= (1+\phi)^2[(1+\phi)z_n - \phi z_{n-1}] - 2\phi(1+\phi)z_n + \phi^2 z_{n-1}\phi + \theta a_n . \end{aligned}$$

Then, $e_n(3) = z_{n+3} - z_n(3) = a_{n+3} + (1+\phi-\theta)a_{n+2} + [(1+\phi)^2 - (1+\phi)\theta - \phi]a_{n+1}$
and $V[e_n(3)] = \sigma^2\{1 + (1+\phi-\theta)^2 + [(1+\phi)^2 - (1+\phi)\theta - \phi]^2\}$.

5.10 (b) (cont'd) There are other quicker ways in which to determine the variance of $e_n(3)$.
The first uses Equation 5.64 on page 239 (as in the solution for the previous model) :
$V[e_n(\ell)] = \sigma^2(1 + \psi_1^2 + \psi_2^2 + \cdots + \psi_{\ell-1}^2)$. Then it is a matter of finding ψ_1 and ψ_2 for the ARIMA(1,1,1) model. This is done by solving the equation
$(1-\phi B)(1-B)(1+\psi_1 B + \psi_2 B^2 + \cdots) = 1 - \theta B$. This results in $\psi_1 = 1 + \phi - \theta$,
$\psi_2 = (1+\phi)\psi_1 - \phi = (1+\phi)(1+\phi-\theta) - \phi$, and in general, $\psi_k = (1+\phi)\psi_{k-1} - \phi\psi_{k-2}$
Alternatively, note that the forecast error satisfies a similar difference equation to the model and forecasts: $e_n(\ell) = (1+\phi)e_n(\ell-1) - \phi e_n(\ell-2) + a_{n+\ell} - \theta a_{n+\ell-1}$.
We always have $e_n(0) = 0$ and $e_n(1) = a_{n+1}$, and we get
$e_n(2) = (1+\phi)e_n(1) - \phi\cdot(0) + a_{n+2} - \theta a_{n+1} = a_{n+2} + (1+\phi-\theta)a_{n+1}$, and then

$$\begin{aligned} e_n(3) &= (1+\phi)e_n(2) - \phi e_n(1) + a_{n+3} - \theta a_{n+2} \\ &= a_{n+3} + (1+\phi-\theta)a_{n+2} + [(1+\phi)(1+\phi-\theta) - \phi]a_{n+1} . \end{aligned}$$

(4) This is an ARIMA(0,2,3) model $z_t = 2z_{t-1} - z_{t-2} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3}$
As always, $e_n(1) = a_{n+1}$. The ψ coefficients are found from
$(1-B)^2(1+\psi_1 B + \psi_2 B^2 + \cdots) = 1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3$, or equivalently,
$(1 - 2B + B^2)(1+\psi_1 B + \psi_2 B^2 + \cdots) = 1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3$, so that
$B:\ \psi_1 - 2 = -\theta_1 \ \rightarrow\ \psi_1 = 2 - \theta_1$,
$B^2:\ \psi_2 - 2\psi_1 + 1 = -\theta_2 \ \rightarrow\ \psi_2 = 2\psi_1 - 1 - \theta_2 = 3 - 2\theta_1 - \theta_2$
Then, $V[e_n(2)] = \sigma^2(1+\psi_1^2) = \sigma^2[1 + (2-\theta)^2]$ and
$V[e_n(3)] = \sigma^2(1+\psi_1^2+\psi_2^2) = \sigma^2[1 + (2-\theta)^2 + (3 - 2\theta_1 - \theta_2)^2]$.
We could have derived in the longer method that was used for the previous model (i.e., find z_{n+3} and $z_n(3)$, and then write $e_n(3) = z_{n+3} - z_n(3)$).

5.11 (a) This is the ARMA(1,1) model, which can be written in autoregressive form as (page 221 of the text) $z_t = \pi_1 z_{t-1} + \pi_2 z_{t-2} + \cdots + a_t$, where $\pi_1 = \phi - \theta$, and $\pi_j = (\phi-\theta)\theta^{j-1}$ for $j \geq 1$. Then, $z_{t+1} = \pi_1 z_t + \pi_2 z_{t-1} + \cdots + a_{t+1}$ so that
$z_t(1) = E(z_{t+1} \mid z_t, z_{t-1}, ...) = \pi_1 z_t + \pi_2 z_{t-1} + \cdots$

$= (\phi - \theta)[z_t + \theta z_{t-1} + \cdots] = (\phi - \theta)\sum_{j=0}^{\infty}\theta^j z_{t-j}$ and

$z_{t+2} = \pi_1 z_{t+1} + \pi_2 z_t + \cdots + a_{t+2}$ so that
$z_t(2) = E(z_{t+1} \mid z_t, z_{t-1}, ...) = \pi_1 z_t(1) + \pi_2 z_t + \pi_3 z_{t-1} + \cdots$

$= (\phi - \theta)^2 \sum_{j=0}^{\infty}\theta^j z_{t-j} + (\phi - \theta)\sum_{j=0}^{\infty}\theta^{j+1} z_{t-j} = (\phi - \theta)\sum_{j=0}^{\infty}\phi\theta^j z_{t-j}$

$= (\pi_1^2 + \pi_2)z_t + (\pi_1\pi_2 + \pi_3)z_{t-1} + \cdots$

(b) The ψ coefficients for the ARMA(1,1) model are $\psi_j = (\phi - \theta)\phi^{k-1}$ for $k \geq 1$.
$V[e_t(\ell)] = \sigma^2(1 + \psi_1^2 + \psi_2^2 + \cdots + \psi_{\ell-1}^2) = \sigma^2[\,1 + (\phi - \theta) + (\phi - \theta)\phi + \cdots + (\phi - \theta)\phi^{\ell-2}]$

$= \sigma^2\left[1 + (\phi - \theta)\left(\frac{1-\phi^{\ell-1}}{1-\phi}\right)\right]$

5.11 (c) As $\phi \to 1$, the forecast $z_n(2) \to (1 - \theta)\sum_{j=0}^{\infty}\theta^j z_{t-j}$, which is the same limit as $z_n(1)$.

As $\phi \to 1$, this model becomes an ARIMA(0,1,1) model.
As $\phi \to 1$, the variance becomes infinite (division by 0 - see part (b) above).

5.12 $z_n(1) = \theta_0 + \phi z_n$ and $z_n(\ell) = \theta_0 + \phi z_n(\ell - 1)$ for $\ell \geq 2$.
We have $z_{100} = 115$, $\widehat{\theta}_0 = 50$ and $\widehat{\phi} = .60$.
Then $\widehat{z}_{100}(1) = 50 + (.6)(115) = 119$, $\widehat{z}_{100}(2) = 50 + (.6)(119) = 121.4$ and
$\widehat{z}_{100}(3) = 50 + (.6)(121.4) = 122.84$.

5.13 For the ARIMA(0,1,1) model, $z_n(\ell) = z_n(1)$ for all $\ell \geq 1$. This $\widehat{z}_{100}(1) = \widehat{z}_{100}(1) = 26$.
We can update the forecasts using Equation 5.85 on page 247 of the text :
$z_{n+1}(\ell) = z_n(\ell + 1) + \psi_\ell[z_{n+1} - z_n(1)]$. In the case of the ARIMA(0,1,1) model,
$\psi_1 = 1 - \theta = .4$, so that $\widehat{z}_{101}(1) = \widehat{z}_{100}(2) + (.4)[z_{101} - \widehat{z}_{100}(1)] = 26 + (.4)(24 - 26) = 25.2$.
Since this is an ARIMA(0,1,1) model, we have $\widehat{z}_{101}(2) = \widehat{z}_{101}(1) = 25.2$.
The 95% prediction interval for the 1-step-ahead forecast from origin t $= 101$ is
$\widehat{z}_{101}(1) \pm u_{.025}\,\sigma = 25.2 \pm 1.96$, and the interval for the 2-step ahead forecast is
$\widehat{z}_{101}(2) \pm u_{.025}\,\sigma\sqrt{1 + (1 - \theta)^2} = 25.2 \pm 1.96\sqrt{1.16}$.

5.14 (a) The ARIMA(1,1,0) model has the form $z_t - z_{t-1} = \phi(z_{t-1} - z_{t-2}) + a_t$, or equivalently,
$z_t = (1 + \phi)z_{t-1} - \phi z_{t-2} + a_t$.
Then $z_{n+1} = (1 + \phi)z_n - \phi z_{n-1} + a_{n+1}$ so that $z_n(1) = (1 + \phi)z_n - \phi z_{n-1}$,
and $z_{n+2} = (1 + \phi)z_{n+1} - \phi z_n + a_{n+2}$ so that
$z_n(2) = (1 + \phi)z_n(1) - \phi z_n = [(1 + \phi)^2 - \phi]z_n - (1 + \phi)\phi z_{n-1}$, and
since $z_{n+\ell} = (1 + \phi)z_{n+\ell-1} - \phi z_{n+\ell-2} + a_{n+\ell}$, it follows that
$z_n(\ell) = (1 + \phi)z_n(\ell - 1) - \phi z_n(\ell - 2)$ for $\ell \geq 3$.
Then $\widehat{z}_{50}(1) = (1.4)(33.9) - (.4)(33.4) = 34.1$, $\widehat{z}_{50}(2) = (1.4)(34.1) - (.4)(33.9) = 34.18$,
$\widehat{z}_{50}(3) = (1.4)(34.18) - (.4)(34.1) = 34.212$, $\widehat{z}_{50}(4) = (1.4)(34.212) - (.4)(34.18) = 34.2248$,
$\widehat{z}_{50}(5) = (1.4)(34.2248) - (.4)(34.212) = 34.22992$.
The ψ coefficients for the ARIMA(1,1,0) model are found from
$(1 - \phi B)(1 - B)(1 + \psi_1 B + \psi_2 B^2 + \cdots) = 1$, so that
$B:\ \psi_1 - 1 - \phi = 0 \ \to\ \psi_1 = 1 + \phi$
$B^2:\ \psi_2 - \psi_1 - \phi\psi_1 + \phi = 0 \ \to \psi_2 = (1 + \phi)\psi_1 - \phi$,
and in general $\psi_j = (1 + \phi)\psi_{j-1} - \phi\psi_{j-2}$.
In this case, $\phi = .40$, so that $\psi_1 = 1.4$, $\psi_2 = (1.4)(1.4) - .4 = 1.56$,
$\psi_3 = (1.4)(1.56) - (.4)(1.4) = 1.624$ and $\psi_4 = (1.4)(1.624) - (.4)(1.56) = 1.6496$
The 95% prediction interval for the ℓ-step-ahead forecast is

$\widehat{z}_n(\ell) \pm 1.96\widehat{\sigma}\sqrt{1+\psi_1^2+\psi_2^2+\cdots+\psi_{\ell-1}^2}$.

5.14 (a) (cont'd) For $\ell = 1$ the interval is $34.1 \pm 1.96(.18) = 34.1 \pm .3528$,
for $\ell = 2$, $34.18 \pm 1.96(.18)\sqrt{1+(1.4)^2} = 34.18 \pm .6070$,
for $\ell = 3$, $34.212 \pm 1.96(.18)\sqrt{1+(1.4)^2+(1.56)^2} = 34.212 \pm .8193$,
for $\ell = 4$, $34.2248 \pm 1.96(.18)\sqrt{1+(1.4)^2+(1.56)^2+(1.624)^2} = 34.2248 \pm .9998$,
for $\ell = 5$, $34.2299 \pm 1.96(.18)\sqrt{1+(1.4)^2+(1.56)^2+(1.624)^2+(1.6496)^2} = 34.2299 \pm 1.1568$.

(b) The updating formula is $\widehat{z}_{n+1}(\ell) = z_n(\ell+1) + \psi_\ell[z_{n+1} - \widehat{z}_n(1)]$. With $z_{51} = 34.2$,
$\widehat{z}_{51}(1) = \widehat{z}_{50}(2) + \psi_1[z_{51} - \widehat{z}_{50}(1)] = 34.18 + (1.4)[34.2 - 34.1] = 34.32$,
$\widehat{z}_{51}(2) = \widehat{z}_{50}(3) + \psi_2[z_{51} - \widehat{z}_{50}(1)] = 34.212 + (1.56)[34.2 - 34.1] = 34.368$,
$\widehat{z}_{51}(3) = \widehat{z}_{50}(4) + \psi_3[z_{51} - \widehat{z}_{50}(1)] = 34.2248 + (1.624)[34.2 - 34.1] = 34.3904$,
$\widehat{z}_{51}(4) = \widehat{z}_{50}(5) + \psi_4[z_{51} - \widehat{z}_{50}(1)] = 34.2299 + (1.6496)[34.2 - 34.1] = 34.3949$.

5.15 For this ARIMA(0,2,2) model we have $z_t = 2z_{t-1} - z_{t-2} + a_t - .82a_{t-1} + .38a_{t-2}$
Then, $z_n(1) = 2z_n - z_{n-1} - .81a_n + .38a_{n-1}$,
$z_n(2) = 2z_n(1) - z_n + .38a_n$ and $z_n(\ell) = 2z_n(\ell-1) - z_n(\ell-2)$ for $\ell \geq 3$.
Following the suggestion given for this problem, we assume that $a_{91} = a_{92} = 0$ and find
$z_{92}(1) = 2z_{92} - z_{91} - .81a_{92} + .38a_{91} = 16.5$.
Then, $a_{93} = z_{93} - z_{92}(1) = 15.9 - 16.5 = -.6$.
Continuing, we get $z_{93}(1) = 2z_{93} - z_{92} - .81a_{93} + .38a_{92} = 16.486$, and
$a_{94} = z_{94} - z_{93}(1) = 15.2 - 16.486 = -1.286$. Successively, we find
$a_{95} = .586$, $a_{96} = .863$, $a_{97} = .476$, $a_{98} = -.742$, $a_{99} = -.182$, $a_{100} = .635$.
Then, $z_{100}(1) = 2z_{100} - z_{99} - .81a_{100} + .38z_{99} = 2(18.2) - 17.3 - (.81)(.635) + (.38)(-.182)$
$= 18.516$.
$z_{100}(2) = 2z_{100}(1) - z_{100} + .38a_{100} = 19.073$, $z_{100}(3) = 2z_{100}(2) - z_{100}(1) = 19.63$
$z_{100}(4) = 2z_{100}(3) - z_{100}(2) = 20.187$, $z_{100}(5) = 20.744$, $z_{100}(6) = 21.301$,
$z_{100}(7) = 21.858$, $z_{100}(8) = 22.415$, $z_{100}(9) = 22.972$, $z_{100}(10) = 23.529$
(subsequent forecasts from $z_{100}(1)$ on increase by .557 over the previous forecast, i.e.,
$z_{100}(\ell) = z_{100}(\ell-1) + .557$ in this example).

5.16 (a) The model can be written in terms of π coefficients in the form
$z_t = \pi_1 z_{t-1} + \pi_2 z_{t-2} + \cdots + a_t$ by solving
$(1-B) = (1-.8B)(1-\pi_1 B - \pi_2 B^2 - \cdots)$.
$B:\ -\pi_1 - .8 = -1 \ \rightarrow\ \pi_1 = .2$
$B^2:\ -\pi_2 + .8\pi_1 = 0 \ \rightarrow\ \pi_2 = .8\pi_1 = (.8)(.2)$
In general, $\pi_j = .8\pi_{j-1} \ \rightarrow\ \pi_j = (.2)(.8)^{j-1}$.

Then $z_t = \sum_{j=1}^{\infty}(.2)(.8)^{j-1}z_{t-j} + a_t$.

The sum of the coefficients (infinite geometric series) is $\sum_{j=1}^{\infty}(.2)(.8)^{j-1} = \frac{.2}{1-.8} = 1$.

(b) $z_t(1) = E(z_{t+1} \mid z_t, z_{t-1}, \ldots) = E(\pi_1 z_t + \pi_2 z_{t-1} + \cdots + a_{t+1}) = \sum_{j=1}^{\infty}(.2)(.8)^{j-1}z_{t-j+1}$

$$z_t(2) = E(z_{t+2} \mid z_t, z_{t-1}, \ldots) = E(\pi_1 z_{t+1} + \pi_2 z_t + \cdots + a_{t+2})$$

$$= \pi_1 z_t(1) + \sum_{j=1}^{\infty}(.2)(.8)^j z_{t-j+1}$$

$$= (.2)\sum_{j=1}^{\infty}(.2)(.8)^{j-1}z_{t-j+1} + \sum_{j=1}^{\infty}(.2)(.8)^j z_{t-j+1}$$

$$= \sum_{j=1}^{\infty}(.2)(.8)^{j-1}z_{t-j+1} = \sum_{j=1}^{\infty}\pi_j^{(2)}z_{t-j+1} ,$$

so that $\pi_j^{(2)} = \pi_j$.

Thus $z_t(2) = z_t(1)$, which is consistent with an ARIMA(0,1,1) model in which $z_t(\ell) = z_t(1)$ for $\ell \geq 1$ (Equation 5.73 on page 244).

(c) $e_t(1) = a_{t+1}$ for any ARIMA time series model. For the ARIMA(0,1,1) model, $e_n(2) = a_{t+2} + (1-\theta)a_{t+1}$. Then $V[e_n(1)] = \sigma^2$, $V[e_n(2)] = \sigma^2[1 + (1-\theta)^2]$ and

$$\begin{aligned}\text{Cov}[e_n(1)\,,e_n(2)] &= E[e_n(1)\,e_n(2)]\\ &= E[a_{t+1}(a_{t+2} + (1-\theta)a_{t+1})]\\ &= E[a_{t+1}a_{t+2}] + E[(1-\theta)a_{t+1}^2)] = 0 + (1-\theta)\sigma^2.\end{aligned}$$

The covariance matrix is

$$\begin{bmatrix} \sigma^2 & (1-\theta)\sigma^2 \\ (1-\theta)\sigma^2 & [1+(1-\theta)^2]\sigma^2 \end{bmatrix} .$$

5.17 Assume that $\ell > j$. Then there will be overlapping a's in $e_t(\ell)$ and $e_{t-j}(\ell)$.

$$\begin{aligned}e_t(\ell) &= a_{t+\ell} + \psi_1 a_{t+\ell-1} + \psi_2 a_{t+\ell-2} + \cdots + \psi_{\ell-1}a_{t+1}\\ &= a_{t+\ell} + \psi_1 a_{t+\ell-1} + \cdots + \psi_j a_{t+\ell-j} + \psi_{j+1}a_{t+\ell-j-1} + \cdots + \psi_{\ell-1}a_{t+1}\\ e_{t-j}(\ell) &= a_{t+\ell-j} + \psi_1 a_{t+\ell-j-1} + \psi_2 a_{t+\ell-j-2} + \cdots + \psi_{\ell-1}a_{t-j+1}\\ &= a_{t+\ell-j} + \psi_1 a_{t+\ell-j-1} + \cdots + \psi_{\ell-j-1}a_{t+1} + \cdots + \psi_{\ell-1}a_{t-j+1}\end{aligned}$$

Since a's at different time points are uncorrelated, the covariance is σ^2 multiplied by the sum of the pairwise products of the coefficients of the overlapping a's. The overlapping a's are
$a_{t+\ell-j}$ which has coefficient ψ_j in $e_t(\ell)$ and coefficient 1 in $e_{t-j}(\ell)$,
$a_{t+\ell-j-1}$ which has coefficient ψ_{j+1} in $e_t(\ell)$ and coefficient ψ_1 in $e_{t-j}(\ell)$
$a_{t+\ell-j-2}$ which has coefficient ψ_{j+2} in $e_t(\ell)$ and coefficient ψ_2 in $e_{t-j}(\ell)$, . . . ,
a_{t+1} which has coefficient $\psi_{\ell-1}$ in $e_t(\ell)$ and coefficient $\psi_{\ell-j-1}$ in $e_{t-j}(\ell)$

$$\begin{aligned}\text{Cov}[e_t(\ell)\,,e_{t-j}(\ell)] &= E[e_t(\ell)e_{t-j}(\ell)]\\ &= \sigma^2[\psi_j + \psi_{j+1}\psi_1 + \psi_{j+2}\psi_2 + \cdots + \psi_{\ell-1}\psi_{\ell-j-1}] = \sigma^2 \sum_{i=j}^{\ell-1}\psi_i\psi_{i-j}.\end{aligned}$$

If $\ell \geq j$ there are no overlapping a's and $e_t(\ell)$ and $e_{t-j}(\ell)$ are uncorrelated .

5.18 $e_t(\ell) = a_{t+\ell} + \psi_1 a_{t+\ell-1} + \psi_2 a_{t+\ell-2} + \cdots + \psi_{\ell-1}a_{t+1}$
$e_t(\ell+j) = a_{t+\ell+j} + \psi_1 a_{t+\ell+j-1} + \cdots + \psi_j a_{t+\ell} + \psi_{j+1}a_{t+\ell-1} + \cdots + \psi_{j+\ell-1}a_{t+1}.$
The over lapping pairs of a's are
$a_{t+\ell}$ which has coefficient 1 in $e_t(\ell)$ and coefficient ψ_j in $e_t(\ell+j)$
$a_{t+\ell-1}$ which has coefficient ψ_1 in $e_t(\ell)$ and coefficient ψ_{j+1} in $e_t(\ell+j)$,
$a_{t+\ell-2}$ which has coefficient ψ_2 in $e_t(\ell)$ and coefficient ψ_{j+2} in $e_t(\ell+j)$, . . . ,
a_{t+1} which has coefficient $\psi_{\ell-1}$ in $e_t(\ell)$ and coefficient $\psi_{j+\ell-1}$ in $e_t(\ell+j)$.

$$\begin{aligned}\text{Cov}[e_t(\ell)\,,e_t(\ell+j)] &= E[e_t(\ell)e_t(\ell+j)]\\ &= \sigma^2[\psi_j + \psi_{j+1}\psi_1 + \psi_{j+2}\psi_2 + \cdots + \psi_{j+\ell-1}\psi_{\ell-1}] = \sigma^2 \sum_{i=0}^{\ell-1}\psi_i\psi_{i+j} .\end{aligned}$$

5.19 $$V[\sum_{\ell=1}^{s}e_t(\ell)] = \sum_{\ell=1}^{s}V[e_t(\ell)] + 2\sum\sum_{1 \leq \ell < m \leq s}\text{Cov}[e_t(\ell), e_t(m)]$$

For $s = 4$,

$$V[\sum_{\ell=1}^{4}e_t(\ell)] = \sum_{\ell=1}^{4}V[e_t(\ell)] + 2[\,\text{Cov}[e_t(1), e_t(2)] + 2\text{Cov}[e_t(1), e_t(3)]$$

$+ 2\text{Cov}[e_t(1), e_t(4)] + 2\text{Cov}[e_t(2), e_t(3)] + 2\text{Cov}[e_t(2), e_t(4)] + 2\text{Cov}[e_t(3), e_t(4)]$
For the ARIMA(0,1,1) model, $\psi_0 = 1$, and $\psi_j = 1 - \theta$ for al $j \geq 1$.
Thus, $V[e_t(1)] = \sigma^2$, $V[e_t(2)] = \sigma^2[1 + (1-\theta)^2]$, $V[e_t(3)] = \sigma^2[1 + 2(1-\theta)^2]$
$V[e_t(4)] = \sigma[1 + 3(1-\theta)^2]$. From Exercise 5.18,
for $\text{Cov}[e_t(1), e_t(2)]$, $\ell = 1$ and $j = 1$, so that $\text{Cov}[e_t(1), e_t(2)] = \sigma^2[\psi_1] = \sigma^2(1-\theta)$,
for $\text{Cov}[e_t(1), e_t(3)]$, $\ell = 1$ and $j = 2$, so that $\text{Cov}[e_t(1), e_t(3)] = \sigma^2[\psi_2] = \sigma^2(1-\theta)$,
$\text{Cov}[e_t(1), e_t(4)] = \sigma^2[\psi_3] = \sigma^2(1-\theta)$,
$\text{Cov}[e_t(2), e_t(3)] = \sigma^2[\psi_1 + \psi_2\psi_1] = \sigma^2[(1-\theta) + (1-\theta)^2]$,
$\text{Cov}[e_t(2), e_t(4)] = \sigma^2[\psi_2 + \psi_3\psi_1] = \sigma^2[(1-\theta) + (1-\theta)^2]$,
$\text{Cov}[e_t(3), e_t(4)] = \sigma^2[\psi_1 + \psi_2\psi_1 + \psi_3\psi_2] = \sigma^2[(1-\theta) + 2(1-\theta)^2]$.

Then $V[\sum_{\ell=1}^{4} e_t(\ell)] = \sigma^2[4 + 12(1-\theta) + 14(1-\theta)^2]$

5.20 (a) ARIMA(0,1,2) . The model is of the form $z_t = z_{t-1} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$.
The forecast function with forecast origin n is
$E(z_{n+1} \mid z_n, ...) = z_n(1) = z_n - \theta_1 a_n - \theta_2 a_{n-1}$,
$z_n(2) = z_n(1) - \theta_2 a_n$, and $z_n(\ell) = z_n(\ell - 1)$ for $\ell \geq 3$.
Value of $z_n(1)$ and $z_n(2)$ are needed, which in turn requires values a_n and a_{n-1} to start this forecasting.
In practice it often is assumed that a_1 and a_2 are both 0, then $z_2(1)$ can be
found, and then $a_3 = z_3 - z_2(1)$ - this process eventually results in values for a_{n-1} and a_n.

(b) ARIMA(1,3,1) . The model is of the form
$z_t - 2z_{t-1} + z_{t-2} = \phi(z_{t-1} - 2z_{t-2} + z_{t-3}) + a_t - \theta a_{t-1}$.
The forecast function with origin n is
$z_n(1) = (2+\phi)z_n - (1+2\phi)z_{n-1} + \phi z_{n-2} - \theta a_n$,
$z_n(2) = (2+\phi)z_n(1) - (1+2\phi)z_n + \phi z_{n-1}$,
$z_n(3) = (2+\phi)z_n(2) - (1+2\phi)z_n(1) + \phi z_n$, and
$z_n(\ell) = (2+\phi)z_n(\ell-1) - (1+2\phi)z_n(\ell-2) + \phi z_n(\ell-3)$ for $\ell \geq 4$.
Forecasting requires the value of a_n to get $z_n(1)$ from which all other forecasts follow
(assuming that z_n and z_{n-1} are known).

5.21 (a) This process is not stationary in the mean, since $E(z_t) = \beta t$ - the z's do not have a common mean.

(b) $w_t = z_t - z_{t-1} = \mu_t - \mu_{t-1} + a_t - a_{t-1} = \epsilon_t + a_t - a_{t-1}$.
$V(w_t) = \sigma_\epsilon^2 + 2\sigma^2$. $\text{Cov}(w_t, w_{t-1}) = \text{Cov}(\epsilon_t + a_t - a_{t-1}, \epsilon_{t-1} + a_{t-1} - a_{t-2})$.
Since the the a's and the ϵ's are uncorrelated, this covariance is $-\sigma^2$ (only $-a_{t-1}$ in w_t
and a_{t-1} in w_{t-1} are correlated). The autocorrelation of the w's at lag 1 is

$\frac{\text{Cov}(w_t, w_{t-1})}{V(w_t)} = \frac{-\sigma^2}{\sigma_\epsilon^2 + 2\sigma^2}$. Since w_t and w_{t-2} have no ϵ's or a's in common, the

covariance and correlation is 0 at lag 2, and the same is true for lag 3 (any lag ≥ 2).

(c) The autocorrelation at lag 1 is non-zero, but the autocorrelation at lag 2 or higher is 0. This suggests an MA(1) model for the w's or equivalently, an ARIMA(0,1,1) model for the z's.

(d) As mentioned on page 244 of the text, forecasting with the ARIMA(0,1,1) model is equivalent to exponential smoothing. These would be MMSE forecasts.

5.22 (a) This is like the model in 5.21 with $\beta = 0$. It is an ARIMA(0,1,1) model with $\theta = 1$.
Then, $z_n(1) = E[z_{n+1} \mid z_n, ...]$

$= E[\mu_{n+1} + a_{n+1} \mid z_n, ...]$
$= E[\mu_{n+1} \mid z_n, ...] = E[\mu_n + \epsilon_{n+1} \mid z_n ...] = \mu_n = z_n - a_n.$

Since $z_n = 100.5$ and $a_n = e_{n-1}(1) = 1$, we have $z_n(1) = 99.5.$

For the ARIMA(0,1,1) model, $z_n(3) = z_n(2) = z_n(1) = 99.5$.

(b) $e_n(1) = z_{n+1} - z_n(1) = \mu_{n+1} + a_{n+1} - \mu_n = \epsilon_{n+1} + a_{n+1}$, which has variance $V(\epsilon_{n+1} + a_{n+1}) = .05 + 1 \ = \ 1.05$ (the ϵ's and the a's are uncorrelated).
$e_n(2) = z_{n+2} - z_n(2) \ = \ \mu_{n+2} + a_{n+2} - \mu_n \ = \ \epsilon_{n+2} + \epsilon_{n+1} + a_{n+2}$, and
$V(\epsilon_{n+2} + \epsilon_{n+1} + a_{n+1}) \ = \ .05 + .05 + 1 \ = \ 1.10$,
$e_n(3) \ = \ z_{n+3} - z_n(3) \ = \ \mu_{n+3} + a_{n+3} - \mu_n \ = \ \epsilon_{n+3} + \epsilon_{n+2} + \epsilon_{n+1} + a_{n+3}$, and
$V(\epsilon_{n+3} + \epsilon_{n+2} + \epsilon_{n+1} + a_{n+1}) = .05 + .05 + .05 + 1 = 1.15.$