Chapter 13 Forecasting with Box-Jenkins Models

We are now at Stage 5 in the forecasting process. Stage 1 deals with stationary time series data for Box-Jenkins Models, this was the subject of Chapter 10. Stage 2 is the Identification Stage of determining the appropriate forecasting model given the data, this was the subject of Chapter 11. Stage 3 is the Estimation Stage of determining the estimated parameters of the Box-Jenkins Model, then Stage 4 is the Diagnostic Checks of the parameter estimates. Stages 3 and 4 were covered in Chapter 12. Now with Chapter 13 we are ready to do some forecasting. We shall walk through several examples of the process, but eventually let the computer take over.

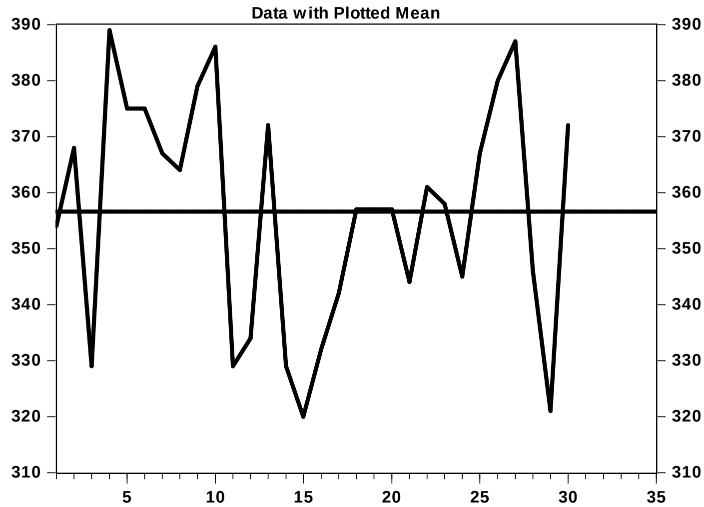
Chapter 13

A Beginning Example

As a beginning example of *ARIMA* model forecasting let us consider an *ARIMA*(0,1,1) model. At the end of Chapter 10 we had shown the equivalence between an *ARIMA*(0,1,1) model and Single Exponential Smoothing. In this chapter we shall use again the data we used for Single Exponential Smoothing in Chapter 4.

Table 13.1

Period	Actual	Period	Actual
1	354	16	332
2	368	17	342
3	329	18	357
4	389	19	357
5	375	20	357
6	375	21	344
7	367	22	361
8	364	23	358
9	379	24	345
10	386	25	367
11	329	26	380
12	334	27	387
13	372	28	346
14	329	29	321
15	320	30	372



Recall that in backshift notation an *ARIMA*(0,1,1) is:

$$(1 - B)Y_t \;=\; (1 \;-\; \theta_1 B)\epsilon_t$$

In expanded notation:

$$Y_t \;=\; Y_{t-1} \;-\; \theta_1\epsilon_{t-1} \;+\; \epsilon_t$$

We estimate θ_1 using computer software, the ϵ_t's are estimated by the fitted residuals, $\hat{\epsilon}_t$'s, determined during the fitting.

$$\hat{Y}_t \;=\; Y_{t-1} \;-\; \hat{\theta}_1\hat{\epsilon}_{t-1} \;+\; \hat{\epsilon}_t$$

In this example we have the parameter estimate by the computer software, $\hat{\theta}_1 =$ 0.9184

$$\hat{Y}_t \;=\; Y_{t-1} \;-\; .9184\hat{\epsilon}_{t-1} \;+\; \hat{\epsilon}_t$$

The Method of Forecasting with Box-Jenkins Models

One-Step Ahead Forecasts

A one-step ahead forecast means that we wish to know expected value of the time series one-step ahead from time period *t*, the latest time period. We wish to know the expected value of Y_{t+1}.

For a one-step ahead forecast the fitted equation is shifted one-step ahead,

$$\hat{Y}_{t+1} \;=\; Y_t \;-\; \hat{\theta}_1\hat{\epsilon}_t \;+\; \hat{\epsilon}_{t+1} \qquad 13.1$$

and then expected values of each term are determined because the one-step ahead forecast is the expected value of the shifted one-step ahead fitted equation.

$$\hat{Y}_t(1) \;=\; E(\hat{Y}_{t+1}) \qquad 13.2$$

Under expectation of both sides of the equation the future values of $\hat{\epsilon}_t$ are 0, so equation 13.1 reduces to

$$E(Y_{t+1}) = Y_t - \hat{\theta}_1 \hat{\epsilon}_t \qquad 13.3$$

Or, $$\hat{Y}_t(1) = Y_t - \hat{\theta}_1 \hat{\epsilon}_t \qquad 13.4$$

The one-step ahead forecast for an *ARIMA*(0,1,1) model

$$\boxed{\hat{Y}_t(1) = Y_t - \hat{\theta}_1 \hat{\epsilon}_t \qquad \mathbf{13.4}}$$

$$\hat{Y}_t(1) = Y_t - .9184\hat{\epsilon}_{tt}$$

Table 13.2

Period t	*Actual* Y_t	*Fit* $\hat{Y}_t$	*Residual* $\hat{\epsilon}_t$	*Forecast* $\hat{Y}_T(\ell)$
1	354	355.6	-1.6	
2	368	355.5	12.5	
3	329	356.5	-27.5	
4	389	354.2	34.8	
⋮	⋮	⋮	⋮	
24	345	353.4	-8.4	
25	367	352.7	14.3	
26	380	353.9	26.1	
27	387	356.0	31.0	
28	346	358.5	-12.5	
29	321	357.5	-36.5	
30	372	354.5	17.5	
31				355.9

Thus, for example, the one-step ahead fit for period 4, $t = 3$,

$$\hat{Y}_3(1) = Y_3 - .9184\hat{\epsilon}_3 = 329 - .9184(\text{-}27.5) = 354.2$$

Or, for example, the one-step ahead fit for period 28, $t = 27$,

$$\hat{Y}_{27}(1) = Y_{27} - .9184\hat{\epsilon}_{27} = 387 - .9184(31.0) = 358.5$$

And lastly, the one-step ahead forecast for period 31, $T = 30$,

$$\hat{Y}_{30}(1) = Y_{30} - .9184\hat{\epsilon}_{30} = 372 - .9184(17.5) = 355.9$$

Two-Step Ahead Forecasts

The two-step ahead forecast for an *ARIMA*(0,1,1) model is

$$\hat{Y}_T(2) = E(Y_{T+2}) \qquad 13.5$$

$$Y_{T+2} = Y_{T+1} - \theta_1 \epsilon_{T+1} + \epsilon_{T+2} \qquad 13.6$$

Under expectation of both sides of equation 13.5 the future terms of $\hat{\epsilon}_T$ reduce to zero,

$$E(Y_{T+2}) = E(Y_{T+1}) \qquad 13.7$$

$\hat{Y}_T(1) = E(Y_{T+1})$, by definition in equation 13.2, thus

The two-step ahead forecast for an *ARIMA*(0,1,1) model.

$$\boldsymbol{\hat{Y}_T(2) = \hat{Y}_T(1)} \qquad \textbf{13.8}$$

Multi-step Ahead Forecasts

It becomes clear that the $\hat{Y}_T(3)$, $\hat{Y}_T(4)$, $\ldots$ are all the same as $\hat{Y}_T(1)$.

$$\hat{Y}_T(1) = Y_T - \hat{\theta}_1 \hat{\epsilon}_T \qquad 8.4$$

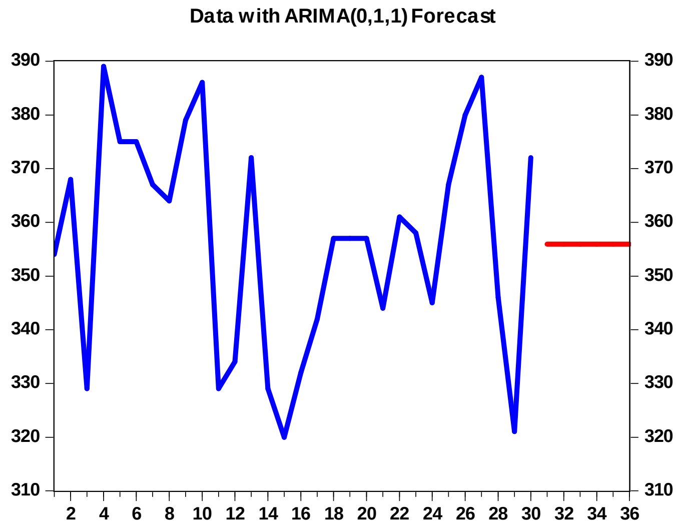
$$\hat{Y}_T(2) = \hat{Y}_T(1)$$

$$\hat{Y}_T(3) = \hat{Y}_T(1)$$

$$\vdots$$

Table 13.3

Period t	*Actual* Y_t	*Fit* $\hat{Y}_t$	*Residual* $\hat{\epsilon}_t$	*Forecast*	
$\vdots$	$\vdots$	$\vdots$	$\vdots$		
24	345	353.4	-8.4		
25	367	352.7	14.3		
26	380	353.9	26.1		
27	387	356.0	31.0		
28	346	358.5	-12.5		
29	321	357.5	-36.5		
30	372	354.5	17.5		
31				355.9	$\hat{Y}_T(1)$
32				355.9	$\hat{Y}_T(2)$
33				355.9	$\hat{Y}_T(3)$
34				355.9	$\hat{Y}_T(4)$
35				355.9	$\hat{Y}_T(5)$
36				355.9	$\hat{Y}_T(6)$



That the one-step ahead forecast is repeated is consistent with the forecasting of single exponential smoothing. Recall in Chapter 4 (page 68) that the last forecast was repeated for single exponential smoothing.

We show again algebraically how an $ARIMA(0,1,1)$ forecast is equivalent to single exponential smoothing.

By definition, the one-step ahead forecast for an $ARIMA(0,1,1)$ is

$$\hat{Y}_t(1) = Y_t - \hat{\theta}_1 \hat{\epsilon}_t \qquad 13.4$$

$\hat{\epsilon}_t$ is the error of fit (residual) at time t,

$$\hat{\epsilon}_t = Y_t - \hat{Y}_{t-1}(1) \qquad 13.9$$

$\hat{Y}_{t-1}(1)$ is the one-step ahead forecast for time t, made at time $t-1$. Substituting (13.9) into (13.4) we have

$$\hat{Y}_t(1) = Y_t - \hat{\theta}_1 (Y_t - \hat{Y}_{t-1}(1)) \qquad 13.10$$

$$\hat{Y}_t(1) = (1 - \hat{\theta}_1) Y_t + \hat{\theta}_1 \hat{Y}_{t-1}(1) \qquad 13.11$$

If we define $\alpha = 1 - \hat{\theta}_1$, then equation (13.11) becomes

$$\hat{Y}_t(1) = \alpha Y_t + (1-\alpha)\hat{Y}_{t-1}(1) \qquad 13.12$$

Equation 13.12 is identical to equation (4.26) [Chapter 4, page 34] when single exponential smoothing was first defined.

Notice in this setting that because

$$\alpha = 1 - \theta_1,$$

we have

$$\alpha = 1 - .9184 = .0816$$

And recall that when testing different choices for α we found that the *SSE* was minimized around .10; (Chapter 8, page 31) well we were pretty close to the “best" alpha.

So the single exponential smoothing model

$$\hat{Y}_t(1) = .0816Y_t + .9184\hat{Y}_{t-1}(1),$$

is equivalent to the *ARIMA* forecasting model

$$\hat{Y}_t(1) = Y_t - .9184\hat{\epsilon}_{tt}$$

In similar fashion it is not difficult to show that an *ARIMA*(0,2,2) is equivalent to *double exponential smoothing*.

These two equivalencies underscore our earlier comment that Box-Jenkins models are more general than exponential smoothing, since exponential smoothing models are just special cases of *ARIMA* models.

Forecasting with Seasonal Models

The above example using an $ARIMA(0,1,1)$ model has no seasonal component. More often than not business/economic data has seasonality, and thus a seasonal model is required. In Chapter 12, an $ARIMA(0,1,1) \times (0,1,1)_{12}$ model was estimated and validated for the Sales Volume data. We now forecast with the model.

$ARIMA(0,1,1) \times (0,1,1)_{12}$

In backshift notation

$$(1 - B)(1 - B^{12})Y_t = (1 - \hat{\theta}_1 B)(1 - \hat{\Theta}_1 B^{12})\hat{\epsilon}_t$$

which expands out to

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - \hat{\theta}_1\hat{\epsilon}_{t-1} - \hat{\Theta}_1\hat{\epsilon}_{t-12} + \hat{\theta}_1\hat{\Theta}_1\hat{\epsilon}_{t-13} + \hat{\epsilon}_t$$

Since $\hat{\theta}_1 = .556$ and $\hat{\Theta}_1 = .6739$, we have

***ARIMA*$(0, 1, 1) \times (0, 1, 1)_{12}$ of the Sales Volume data**

$$\boldsymbol{Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} - .556\hat{\epsilon}_{t-1} - .674\hat{\epsilon}_{t-12} + .375\hat{\epsilon}_{t-13} + \hat{\epsilon}_t} \qquad \textbf{13.13}$$

One-Step Ahead Forecast

By definition a one-step ahead forecast is

$$\hat{Y}_T(1) = \mathrm{E}(\mathrm{Y}_{T+1})$$

Equation 13.13 is shifted ahead one period

$$Y_{T+1} = Y_T + Y_{T-11} - Y_{T-12} - .556\hat{\epsilon}_T - .674\hat{\epsilon}_{T-11} + .375\hat{\epsilon}_{T-12} + \hat{\epsilon}_{T+1} \qquad 13.14$$

and then we take expectation of both sides of equation 13.14.

$$E(\mathrm{Y}_{T+1}) = E(\mathrm{Y}_T + Y_{T-11} - Y_{T-12} - .556\hat{\epsilon}_T - .674\hat{\epsilon}_{T-11} + .375\hat{\epsilon}_{T-12} + \hat{\epsilon}_{T+1}) \qquad 13.15$$

All terms in equation 13.15 are known values so they unaffected by expectation, with the exception that $E(\hat{\epsilon}_{T+1}) = 0$.

Thus, equation 13.15 reduces to

$$\hat{Y}_T(1) = Y_T + Y_{T-11} - Y_{T-12} - .556\hat{\epsilon}_T - .674\hat{\epsilon}_{T-11} + .375\hat{\epsilon}_{T-12} \qquad 13.16$$

In this example using the Sales Volume data we have $T = 48$, so a one-step ahead forecast is a forecast for period 49,

$$\hat{Y}_{48}(1) = Y_{48} + Y_{37} - Y_{36} - .556\hat{\epsilon}_{48} - .674\hat{\epsilon}_{37} + .375\hat{\epsilon}_{36} \qquad 13.17$$

We list below the observations from Period 35 and the corresponding residuals.

Table 13.4

Period	*Data*	*Fit*	*Residual*	*Forecast*
⋮	⋮	⋮	⋮	
35	490.5	469.22	21.48	
36	532.5	526.02	6.48	
37	440.0	445.62	-5.62	
38	472.0	456.95	15.05	
39	494.5	465.50	29.00	
40	512.0	493.17	18.83	
41	519.0	460.50	58.50	

42	526.0	542.21	-16.21	
43	530.5	545.59	-15.09	
44	5513.0	548.74	9.26	
45	557.5	573.15	-15.65	
46	590.0	584.41	5.59	
47	636.4	634.07	2.33	
48	677.4	680.72	-3.32	
49				592.96

To forecast we substitute the appropriate Data and Residual values into equation 13.17.

$$\hat{Y}_{48}(1) = Y_{48} + Y_{37} - Y_{36} - .556\hat{\epsilon}_{48} - .674\hat{\epsilon}_{37} + .375\hat{\epsilon}_{36}$$

$$\hat{Y}_{48}(1) = 677.4 + 440.0 - 532.5 - .556(\text{-}3.32) - .674(\text{-}5.62) + .375(6.48)$$

$$\hat{Y}_{48}(1) = 592.96$$

Forecasting with Seasonal Models continued:
Two-Step Ahead Forecast

A two-step ahead forecast involves shifting equation 13.20 two periods ahead and taking expectations.

$$\hat{Y}_T(2) = E(\mathrm{Y}_{T+2})$$

$$\hat{Y}_T = Y_{T-1} + Y_{T-12} - Y_{T-13} - .556\hat{\epsilon}_{T-1} - .674\hat{\epsilon}_{T-12} + .375\hat{\epsilon}_{T-13} + \hat{\epsilon}_T \qquad 13.20$$

$$Y_{T+2} = Y_{T+1} + Y_{T-10} - Y_{T-11} - .556\hat{\epsilon}_{T+1} - .674\hat{\epsilon}_{T-10} + .375\hat{\epsilon}_{T-11} + \hat{\epsilon}_{T+2}$$

$$E(\mathrm{Y}_{T+2}) = E(\mathrm{Y}_{T+1} + Y_{T-10} - Y_{T-11} - .556\hat{\epsilon}_{T+1} - .674\hat{\epsilon}_{T-10} + .375\hat{\epsilon}_{T-11} + \hat{\epsilon}_{T+2})$$

$$E(\mathrm{Y}_{T+2}) = E(\mathrm{Y}_{T+1}) + E(\mathrm{Y}_{T-10}) - E(\mathrm{Y}_{T-11}) - .556\mathrm{E}(\hat{\epsilon}_{T+1}) - .674\mathrm{E}(\hat{\epsilon}_{T-10}) + .375\mathrm{E}(\hat{\epsilon}_{T-11}) + E(\hat{\epsilon}_{T+2}) \qquad 13.18$$

Notice in equation 13.18 that the terms containing $\hat{\epsilon}_{T+1}$ and $\hat{\epsilon}_{T+2}$ both go to zero under expectation. Notice too that $E(\mathrm{Y}_{T+1}) = \hat{Y}_T(1)$, so that $\hat{Y}_T(2)$ is recursively constructed using the previous forecast.

Equation 13.18 reduces to

$$\hat{Y}_T(2) = \hat{Y}_T(1) + Y_{T-10} - Y_{T-11} - .674(\hat{\epsilon}_{T-10}) + .375(\hat{\epsilon}_{T-11}) \qquad 13.19$$

Again, since $T = 48$ equation 13.19 becomes

$$\hat{Y}_{48}(2) = \hat{Y}_{48}(1) + Y_{38} - Y_{37} - .674(\hat{\epsilon}_{38}) + .375(\hat{\epsilon}_{37}) \qquad 13.20$$

Substituting in the appropriate values from Table 13.1, we determine the two-step ahead forecast

$$\hat{Y}_{48}(2) = 592.96 + 472.0 - 440.0 - .674(15.05) + .375(\text{-}5.62) = 612.714$$

Table 13.5

Period	*Data*	*Fit*	*Residual*	*Forecast*
⋮	⋮	⋮	⋮	
35	490.5	469.22	21.48	
36	532.5	526.02	6.48	
37	440.0	445.62	-5.62	
38	472.0	456.95	15.05	

39	494.5	465.50	29.00	
40	512.0	493.17	18.83	
41	519.0	460.50	58.50	
42	526.0	542.21	-16.21	
43	530.5	545.59	-15.09	
44	558.0	548.74	9.26	
45	557.5	573.15	-15.65	
46	590.0	584.41	5.59	
47	636.4	634.07	2.33	
48	677.4	680.72	-3.32	
49				592.96
50				612.71

Forecasting with Seasonal Models Continued:
Multi-Period Forecasts

This process is repeated for multi-period forecasts. Notice that we can think of the forecasts for leads up to 13 periods as a combination of previous forecasts and historical data and historical residuals. Beyond the lead of 13 periods, i.e. when $\ell > 13$, the forecasts may be thought of as being made up of as a combination of previous forecasts.

$$\hat{Y}_{48}(1) = Y_{48} + Y_{37} - Y_{36} - .556\hat{e}_{48} - .674\hat{e}_{37} + .375\hat{e}_{36} = 592.6$$

$$\hat{Y}_{48}(2) = \hat{Y}_{48}(1) + Y_{38} - Y_{37} - .674\hat{e}_{38} + .375\hat{e}_{37} = 612.71$$

$$\hat{Y}_{48}(3) = \hat{Y}_{48}(2) + Y_{39} - Y_{38} - .674\hat{e}_{39} + .375\hat{e}_{38} = 621.32$$

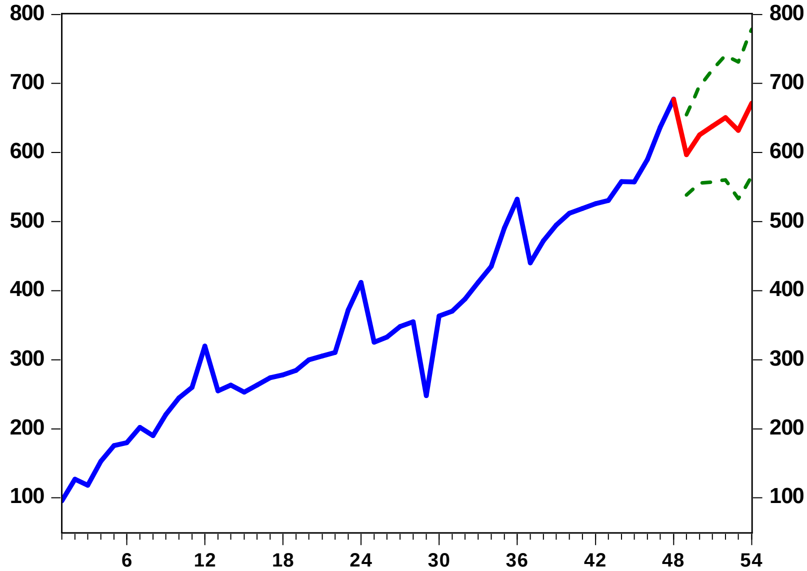
$$\vdots$$

$$\hat{Y}_{48}(6) = \hat{Y}_{48}(5) + Y_{42} - Y_{41} - .674\hat{e}_{42} + .375\hat{e}_{41} = 651.48$$

Table 13.6

Period	*Data*	*Fit*	*Residual*	*Forecast*
⋮	⋮	⋮	⋮	
35	490.5	469.22	21.48	
36	532.5	526.02	6.48	
37	440.0	445.62	-5.62	
38	472.0	456.95	15.05	
39	494.5	465.50	29.00	
40	512.0	493.17	18.83	
41	519.0	460.50	58.50	
42	526.0	542.21	-16.21	
43	530.5	545.59	-15.09	
44	558.0	548.74	9.26	
45	557.5	573.15	-15.65	
46	590.0	584.41	5.59	
47	636.4	634.07	2.33	
48	677.4	680.72	-3.32	
49				592.96
50				612.71
51				621.32
52				637.00
53				611.64
54				651.48

The graph below presents the forecasts and confidence intervals of forecast Notice how the *ARIMA* model captures the seasonality of the data. It drops down for period 49 but climbs for periods 50 through 54.



A Second Example of Forecasting Seasonal Data

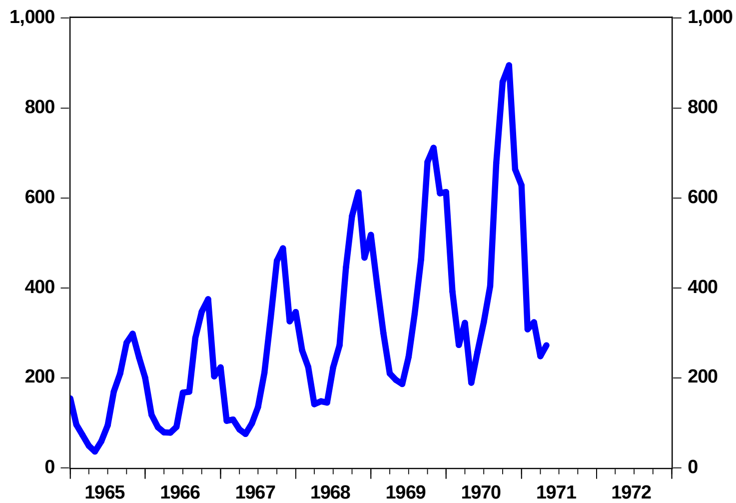
Stage 1 ***Data Collection and Analysis***

Let us consider the Sales Data found in Appendix: Table A.1.

Table 13.7

Sales Data

	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1965	154	96	73	49	36	59	95	169	210	278	298	245
1966	200	118	90	79	78	91	167	169	289	347	375	203
1967	223	104	107	85	75	99	135	211	335	460	488	326
1968	346	261	224	141	148	145	223	272	445	560	612	467
1969	518	404	300	210	196	186	247	343	464	680	711	610
1970	613	392	273	322	189	257	324	404	677	858	895	664
1971	628	308	324	248	272							



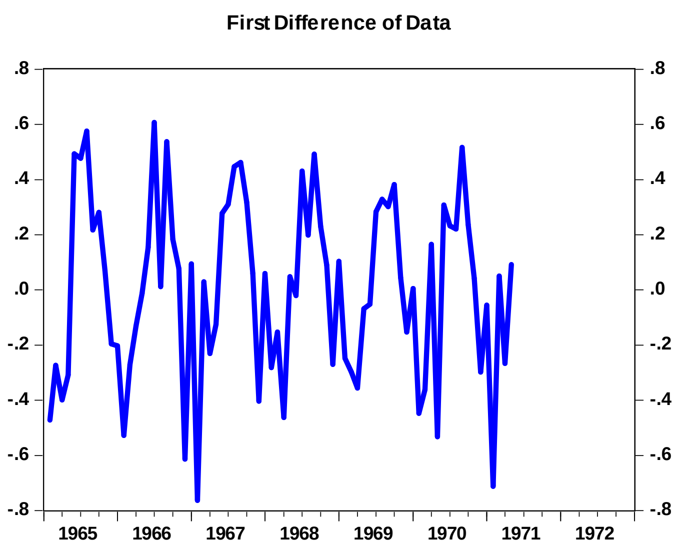
Stage 2 **Model Identification**

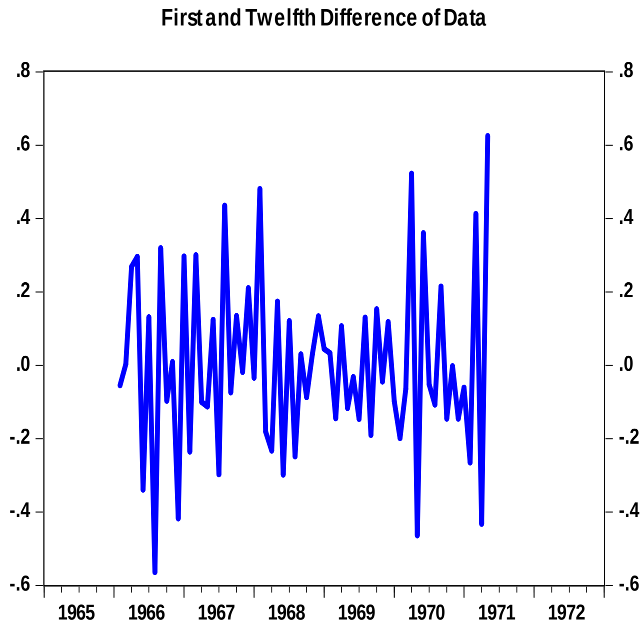
The data clearly has trend, so before attempting to identify the particular model we must difference the data. Since it is monthly data we take both the first and twelfth difference of the data.

Table 13.8

Obs		Data	First	First and Twelfth Difference
1	JAN65	154	*	*
2		96	- 58	*
3		73	- 23	*
4		49	- 24	*
5		36	- 13	*
6		59	23	*

7	JUL65	95	36	*
8		169	74	*
9		210	41	*
10		278	68	*
11		298	20	*
12		245	- 53	*
13	JAN66	200	- 45	*
14		118	- 82	-24
15		90	- 28	-5
16		79	- 11	13
17		78	-1	12
18		91	13	-10
19	JUL66	167	76	40
20		169	2	-72
21		289	120	79
⋮		⋮	⋮	⋮
73	JAN71	628	- 36	-39
74		308	- 320	-99
75		324	161	35
76		248	- 76	-125
77		272	241	57





The data now appears stationary so we request an *ACF* and a *PACF* of the ***differenced data***.

Table 13.9

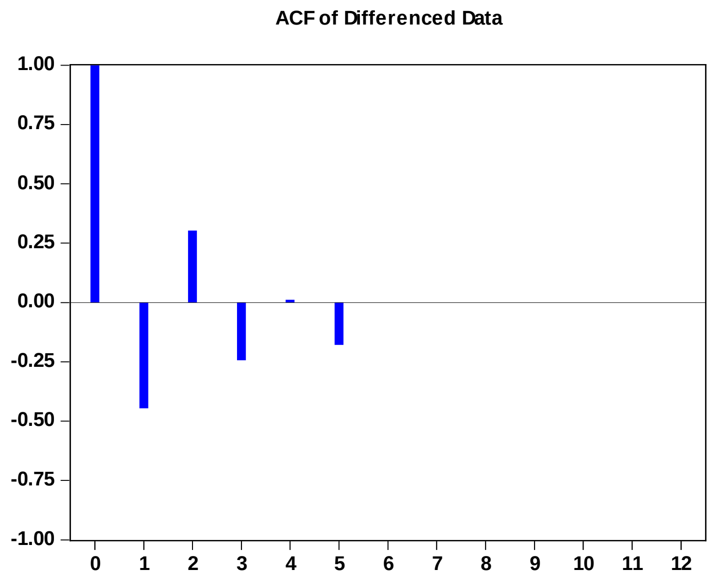
ACF

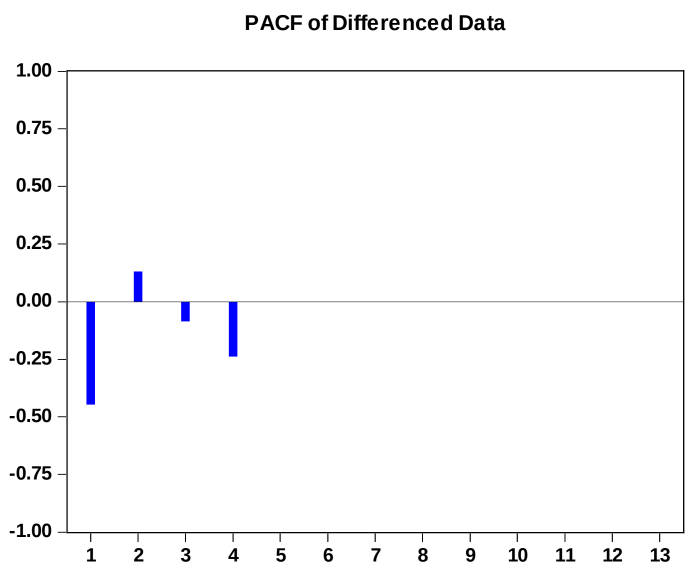
Lags	0	1	2	3	4	5	. . .
Autocorrelation	1	-0.445	0.303	-0.242	0.012	-0.177	

Table 13.10

PACF

Lags	1	2	3	4	5	. . .
Partial Autocorrelation	-0.445	0.131	-0.085	-0.187	-0.238	





It is not the clearest identification, but it appears to be an $AR(1)$ regular, an $MA(2)$ regular and an $MA(1)$ seasonal. So we request of the computer software to estimate an $ARIMA(1,1,2) \times (0,1,1)_{12}$.

Table 13.11

Estimates at each iteration

Iteration	SSE	Parameters			
0	253653	0.100	0.100	0.100	0.100
1	243232	0.222	0.250	0.082	0.107
2	237651	0.359	0.400	0.067	0.110
3	232884	0.498	0.550	0.051	0.114

⋮	⋮	⋮	⋮	⋮	⋮

Final Estimates of Parameters

Type	Estimate	St. Dev	t-ratio
AR 1	0.5727	0.1528	3.75
MA 1	0.6738	0.1706	5.67
MA 2	-0.0320	0.1750	-0.18
SMA 12	0.3066	0.1998	1.53

The *t-ratios* of the *regular* $MA(2)$ and *seasonal* $MA(1)$ are not satisfactory. This suggests that we have not identified the correct model. Thus we re-estimate the model by reducing the *regular* $MA(2)$ to a *regular* $MA(1)$. So we request of the computer software to estimate an $ARIMA(1,1,1) \times (0,1,1)_{12}$.

Table 13.12

Estimates at each iteration

Iteration	SSE	Parameters		
0	239164	0.100	0.100	0.100
1	204948	-0.050	0.247	0.143
2	201069	-0.200	0.117	0.154
3	196940	-0.350	-0.024	0.164
⋮	⋮	⋮	⋮	⋮
18	187170	-0.820	-0.473	0.137

Final Estimates of Parameters

Type	Estimate	St. Dev	t-ratio
AR 1	-0.8202	0.1779	-4.61
MA 1	-0.4732	0.2618	-1.81
SMA 12	0.1370	0.1966	0.70

The t-ratios have worsened, indicating the incorrect identification

As a last resort we have the software estimate the old standby model of an $ARIMA(0,1,1) \times (0,1,1)_{12}$.

Table 13.13

Estimates at each iteration

Iteration	SSE	Parameters	
0	222920	0.100	0.100
1	207985	0.250	0.130
2	204302	0.339	0.195
3	204225	0.321	0.229
⋮	⋮	⋮	⋮

Final Estimates of Parameters

Type	Estimate	St. Dev	t-ratio
MA 1	0.2914	0.1564	1.86
SMA 12	0.2789	0.2044	1.36

This $ARIMA(0,1,1) \times (0,1,1)_{12}$ model has poor *t-ratios* and other poor diagnostic statistics. This too is apparently not the correct model identification.

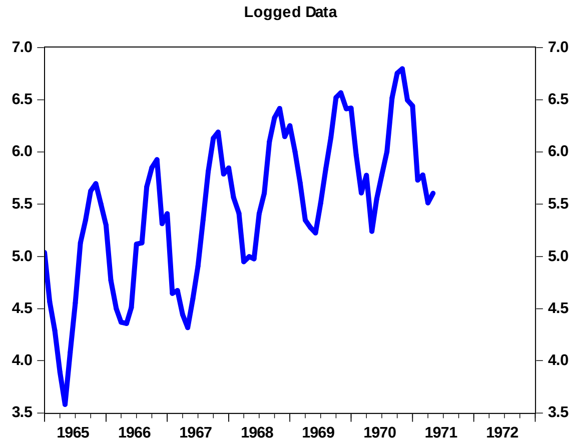
Forecasting Non-stationary Data

The source of the problem lies in the fact that the data is not yet stationary. Returning to the orginal graph of the data, Figure 13-3, we observe that ***the variance of the data is not constant***. As the data grows over time the variance also grows. To remove growing variance so that the data will be stationary we take the natural logarithm of the data.

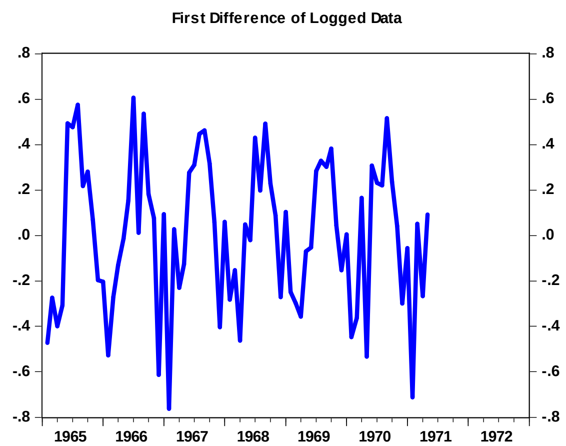
Table 13.14

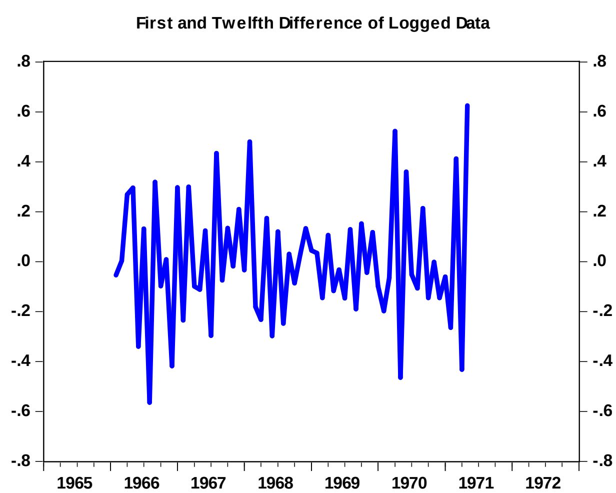
Obs		Original Data	Natural Log of the Data	First Difference	First and Twelfth Difference
1	JAN65	154	5.037	*	*
2		96	4.564	-0.473	*
3		73	4.290	-0.274	*
4		49	3.892	-0.399	*
5		36	3.584	-0.308	*
6		59	4.078	0.494	*
7	JUL65	95	4.554	0.476	*
8		169	5.130	0.576	*
9		210	5.347	0.217	*
10		278	5.628	0.281	*
11		298	5.697	0.069	*
12		245	5.501	-0.196	*
13	JAN66	200	5.298	-0.203	*
14		118	4.771	-0.528	-0.055
15		90	4.500	-0.271	0.003
16		79	4.369	-0.130	0.268
17		78	4.357	-0.013	0.296
18		91	4.511	0.154	-0.340
19	JUL66	167	5.118	0.607	0.131
20		169	5.130	0.012	-0.564
21		289	5.666	0.537	0.319
⋮		⋮	⋮	⋮	⋮
73	JAN71	628	6.443	-0.056	-0.061
74		308	5.730	-0.712	-0.265
75		324	5.781	0.051	0.412
76		248	5.513	-0.267	-0.432
77		272	5.606	0.092	0.625

Notice now in the graph of the logged data, that the variance is no longer growing over time.

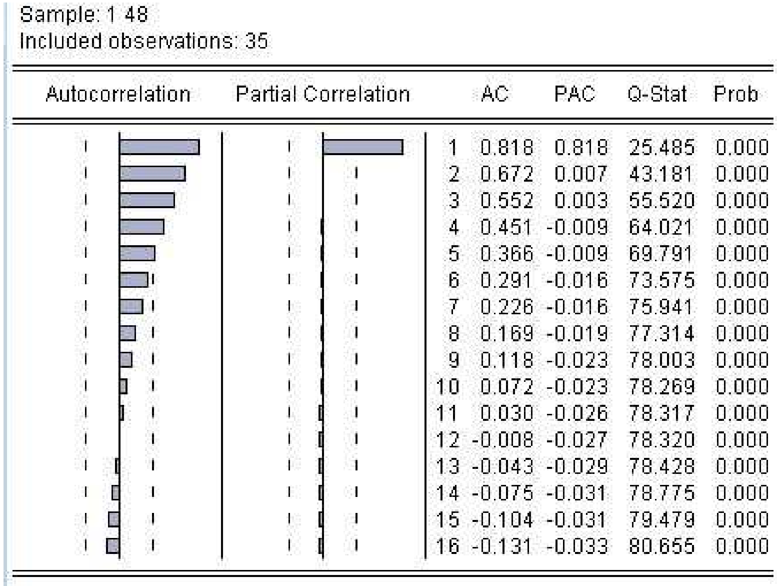


We then graph the first and twelfth difference of the logged data to observe stationarity.





We request of the computer software the *ACF* and *PACF*. We request of the software that the *ACF* and *PACF* be taken to 25 lags in order to check for *seasonal* terms.


Sample: 1 48
Included observations: 35

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.818	0.818	25.485	0.000
		2	0.672	0.007	43.181	0.000
		3	0.552	0.003	55.520	0.000
		4	0.451	-0.009	64.021	0.000
		5	0.366	-0.009	69.791	0.000
		6	0.291	-0.016	73.575	0.000
		7	0.226	-0.016	75.941	0.000
		8	0.169	-0.019	77.314	0.000
		9	0.118	-0.023	78.003	0.000
		10	0.072	-0.023	78.269	0.000
		11	0.030	-0.026	78.317	0.000
		12	-0.008	-0.027	78.320	0.000
		13	-0.043	-0.029	78.428	0.000
		14	-0.075	-0.031	78.775	0.000
		15	-0.104	-0.031	79.479	0.000
		16	-0.131	-0.033	80.655	0.000

The original series had 77 observations, after differencing 64 observations. Testing for significant spikes we use:

$$\frac{\pm 2}{\sqrt{n}} = \frac{\pm 2}{\sqrt{64}} = \pm .25$$

It appears by the *ACF* and *PACF* that the logged Sales Data is an *ARIMA*$(1,1,0) \times (0,1,1)_{12}$. So we request of the computer software to estimate such a model on the original *logged* Sales Data.

Stage 3 Estimates of the Model Parameters

Table 13.15

Estimates at each iteration

Iteration	SSE	Parameters	
0	3.99682	0.100	0.100
1	3.29109	-0.050	0.136
2	2.74576	-0.200	0.187
3	2.33078	-0.350	0.263
⋮	⋮	⋮	⋮

Final Estimates of Parameters

Type	Estimate	St. Dev	t-ratio
AR 1	-0.4580	0.1144	-4.01
SMA 12	0.7954	0.1165	6.83

The $ARIMA(1,1,0) \times (0,1,1,)_{12}$ model

In Backshift notation:

$$(1 - \hat{\phi}_1 B)(1 - B)(1 - B^{12}) lnY_t = (1 - \hat{\Theta}_1 B^{12})\hat{\epsilon}_t$$

In Expanded notation:

$$lnY_t = (1 + \hat{\phi}_1) lnY_{t-1} - \hat{\phi}_1 lnY_{t-2} - (1 + \hat{\phi}_1) lnY_{t-13} + \hat{\phi}_1 lnY_{t-14} - \hat{\Theta}_1 \hat{\epsilon}_{t-12} + \hat{\epsilon}_t$$

With $\hat{\phi}_1 = \text{-0.4580}, \hat{\Theta}_1 = .7954,$

$$lnY_t = (1 + (\text{-0.458})) lnY_{t-1} - (\text{-0.458}) lnY_{t-2} - (1 + (\text{-0.458})) lnY_{t-13} + (\text{-0.458}) lnY_{t-14} - (0.795)\hat{\epsilon}_{t-12} + \hat{\epsilon}_t$$

$$lnY_t = 0.542 lnY_{t-1} + 0.458 lnY_{t-2} - 0.542 lnY_{t-13} - 0.458 lnY_{t-14} - 0.795\hat{\epsilon}_{t-12} + \hat{\epsilon}_t$$

Stage 4 Diagnostics and Residual Analysis

Diagnostics

Table 13.16

Estimates at each iteration

Iterat ion	SSE	Parameters	
0	3.99682	0.100	0.100
1	3.29109	-0.050	0.136
2	2.74576	-0.200	0.187
3	2.33078	-0.350	0.263
⋮	⋮	⋮	⋮
11	1.79772	-0.458	0.795

Final Estimates of Parameters

Type	Estimate	St. Dev	t-ratio
AR 1	-0.4580	0.1144	-4.01
SMA 12	0.7954	0.1165	6.83

Differencing: 1 regular, 1 seasonal of order 12

No. of obs.: Original series 77, after differencing 64

Residuals: SS = 1.42235 (backforecasts excluded)

MS = 0.02294 DF = 62

Modified Box-Pierce chisquare statistic

Lag	12		24		36		48	
Chisquare	20.2	(DF = 10)	30.0	(DF = 22)	47.1	(DF = 34)	66.5	(DF = 46)

The model is estimated in 11 iterations of the non-linear grid search – this is good.

The Back Forecast residuals die out in lag – this is also good.

The *t-ratios* are both good for the $AR(1)$ estimate; $\hat{\phi}_1 = -0.4580$, *t-ratio* of -4.01 and the $SMA(1)$ estimate; $\hat{\Theta}_1 = .7954$, *t-ratio* of 6.83. These too are good.

The *Modified Box-Pierce* χ^2 *statistic* at lag 12 (d.f. = 10) is 20.2. At 5% level of significance χ^2 is 18.3, compared to 20.2. Not quite random residuals at the first 12 lags. At lag 24 (d.f. = 22) it is 30.0, compared with 33.9 – this is good. At lag 36 (d.f. = 34) it is 47.1, compared with about 48.6 – this is good. At lag 48 (d.f. = 46) it is 66.5, compared with 62.8 – oh well, can't win them all.

The low correlation between the estimates is .221 – this is good.

This appears to be reasonable identification; that the logged sales data is an *ARIMA*$(1,1,0) \times (0,1,1,)_{12}$ process.

For comparison purposes we request of the computer software to estimate the old standby of an *ARIMA*$(0,1,1) \times (0,1,1)_{12}$ model.

Table 13.17

Estimates at each iteration

Iteration	SSE	Parameters	
0	3.21012	0.100	0.100
1	2.75609	0.250	0.158
2	2.40314	0.400	0.269
3	2.29972	0.440	0.419
⋮	⋮	⋮	⋮

Final Estimates of Parameters

Type	Estimate	St. Dev	t-ratio
MA 1	0.4626	0.1205	3.84
SMA 12	0.5557	0.1208	4.60

Differencing:	1 regular, 1 seasonal of order 12		
No. of obs.:	Original series 77, after differencing 64		
Residuals:	SS =	2.22320	(backforecasts excluded)
	MS =	0.03586	DF = 62

Modified Box-Pierce chisquare statistic

Lag		12		24		36		48
Chisquare	25.4	(DF = 10)	30.8	(DF = 22)	36.3	(DF = 34)	64.7	(DF = 46)

The two models are close. The second requires fewer iterations and again has back forecasts that die out. The *t-ratios* of the second model are not quite as good. Notice also that the *MSE* of the second model is 0.03586, slightly larger than that the *MSE* of the first model, 0.02294. The second model's χ^2 statistics are slightly worse, and the correlation between estimates is somewhat larger.

Let's stay with the first model.

Residual Analysis

The residuals are differences between the actuals and the fits.

Table 13.18

Natural Log

Obs		Data	Fitted Value	Residual
1	JAN65	5.037	*	*
2		4.564	*	*
3		4.290	*	*
4		3.892	*	*
5		3.584	*	*
6		4.078	*	*
7	JUL65	4.554	*	*
8		5.130	*	*
9		5.347	*	*
10		5.628	*	*
11		5.697	*	*
12		5.501	*	*
13	JAN66	5.298	*	*
14		4.771	4.912	-0.142
15		4.500	4.546	-0.047
16		4.369	4.285	0.084
17		4.357	4.104	0.252
18		4.511	4.473	0.038
19	JUL66	5.118	4.948	0.170
20		5.130	5.408	-0.278
21		5.666	5.632	0.035
22		5.849	5.873	-0.024
⋮		⋮	⋮	⋮
64		5.775	5.398	0.377
65		5.242	5.457	-0.216
66		5.549	5.557	-0.008
67	JUL70	5.781	5.860	-0.079
68		6.001	6.166	-0.165
69		6.518	6.433	0.085
70		6.755	6.756	-0.001
71		6.797	6.846	-0.049
72		6.498	6.510	-0.012
73	JAN71	6.443	6.516	-0.074
74		5.730	6.034	-0.304
75		5.781	5.614	0.167
76		5.513	5.457	0.056
77		5.606	5.350	0.256

- **Diagnostic Plots**

Plotting the Data and the Fitted Values we observe how the $ARIMA(1,1,0) \times (0,1,1)_{12}$ tracks the data.

Figure 13-1

We then plot the Residuals over the same time period.

Figure 13-2

Figure 13-3

From the plot, the residuals appear to be random with no pattern.
In Table 13.12 above we are given information about the residuals.

Residuals:	SS = 1.42235	(backforecasts excluded)
	MS = 0.02294	DF = 62

In our notation,

$$SSE = \sum \hat{\epsilon}^2 = 1.42235$$

$$\sigma_\epsilon^2 = MSE = \frac{\sum \hat{\epsilon}^2}{64-2} = \frac{1.42235}{62} = 0.02294$$

$$\sigma_\epsilon = \sqrt{0.02294} = 0.15146$$

The two standard errors of the model, $2\sigma_\epsilon$, on either side of 0 should capture about 95% of the residuals.

$$\pm\ 2\sigma_\epsilon = \pm\ 0.303$$

Hence, we draw lines at ± 0.303 and expect no more than about 5% of the residuals to fall outside the lines. Indeed, only two residuals are greater in absolute value than 0.303, and only by a small amount. Consequently, no real outliers of residuals.

We also plot the residuals against the log of Sales data to check for a pattern in the residuals. There is no apparent pattern in the residual plot.

Figure 13-4

• ACF of the Residuals

Another diagnostic check of the residuals is to determine the ACF of the residuals. If the residuals are purely random, like “white noise," then the *ACF* would identify an *ARIMA*(0, 0, 0) model.

Figure 13-5

Figure 13-17 of the *ACF* of the residuals essential supports are view that the residuals are white noise. There is one odd significant spike at lag 11 but no real recognizable pattern in the *ACF*.

Stage 5 Forecasting and Confidence Intervals

All the Diagnostic checks of the identified *ARIMA* model are satisfactory so we proceed to forecast.

The $ARIMA(1,1,0) \times (0,1,1,)_{12}$ model written in backshift notation is:

$$(1 - \hat{\phi}_1 B)(1 - B)(1 - B^{12}) lnY_t = (1 - \hat{\Theta}_1 B^{12})\hat{\epsilon}_t$$

In expanded notation:

$$lnY_t = (1 + \hat{\phi}_1)lnY_{t-1} - \hat{\phi}_1 lnY_{t-2} + lnY_{t-12} - (1 + \hat{\phi}_1)lnY_{t-13} + \hat{\phi}_1 lnY_{t-14} - \hat{\Theta}_1\hat{\epsilon}_{t-12} + \hat{\epsilon}_t$$

With $\hat{\phi}_1 = -0.4580$, $\hat{\Theta}_1 = .7954$, the full model is

$$lnY_t = \left(1 + (-0.458)\right)lnY_{t-1} - (-0.458)lnY_{t-2} + lnY_{t-12} - \left(1 + (-0.458)\right)lnY_{t-13} + (-0.458)lnY_{t-14} - (0.795)\hat{\epsilon}_{t-12} + \hat{\epsilon}_t$$

The model is thus:

$$lnY_t = 0.542lnY_{t-1} + 0.458lnY_{t-2} + lnY_{t-12} - 0.542lnY_{t-13} - 0.458lnY_{t-14} - 0.795\hat{\epsilon}_{t-12} + \hat{\epsilon}_t \qquad 13.21$$

Forecasting

One-Step Ahead Forecast

A ***One-Step Ahead Forecast*** is determined by shifting all terms one-step ahead in equation 13.21:

$$ln\hat{Y}_T(1) = 0.542\ln\mathrm{Y}_T + 0.458lnY_{T-1} + lnY_{T-11} - 0.542lnY_{T-12} - 0.458lnY_{T-13} - 0.795\hat{\epsilon}_{T-11}$$ [1]

In this example, because $T = 77$, a One-Step Ahead forecast is the forecast for period 78:

$$ln\hat{Y}_{77}(1) = 0.542lnY_{77} + 0.458lnY_{76} + lnY_{66} - 0.542lnY_{65} - 0.458lnY_{64} - 0.795\hat{\epsilon}_{66}$$

Table 13.19

Obs		Natural Log Data	Fitted Value	Residual
		⋮	⋮	⋮
64		5.775	5.398	0.377
65		5.242	5.457	-0.216
66		5.549	5.557	-0.008
67	JUL70	5.781	5.860	-0.079
68		6.001	6.166	-0.165
69		6.518	6.433	0.085
70		6.755	6.756	-0.001
71		6.797	6.846	-0.049
72		6.498	6.510	-0.012
73	JAN71	6.443	6.516	-0.074
74		5.730	6.034	-0.304
75		5.781	5.614	0.167
76		5.513	5.457	0.056
77		5.606	5.350	0.256

We substitute in the appropriate values from Table 13.15 above to determined the forecast for period 78.

1

$$ln\hat{Y}_{77}(1) = 0.542lnY_{77} + 0.458lnY_{76} + lnY_{66} - 0.542lnY_{65} - 0.458lnY_{64} - 0.795\hat{\epsilon}_{66}$$

$$ln\hat{Y}_{77}(1) = 0.542(5.606) + 0.458(5.513) + 5.549 - 0.542(5.242) - 0.458(5.775) - 0.795(-0.008)$$

$ln\hat{Y}_{77}(1) = 5.633$; the One-Step Ahead Forecast for period 78.

Two-Step Ahead Forecast

A ***Two-Step Ahead Forecast*** is determined by shifting all terms two-steps ahead in equation 13.21:

$$lnY_T = 0.542lnY_{T-1} + 0.458lnY_{T-2} + lnY_{T-12} - 0.542lnY_{T-13} - 0.458lnY_{T-14} - 0.795\hat{\epsilon}_{T-12} + \hat{\epsilon}_T$$

$$lnY_{T+2} = 0.542lnY_{T+1} + 0.458lnY_T + lnY_{T-10} - 0.542lnY_{T-11} - 0.458lnY_{T-12} - 0.795\hat{\epsilon}_{T-10} + \hat{\epsilon}_{T+2}$$

Under expectation lnY_{T+2} and lnY_{T+1} become $ln\hat{Y}_T(2)$ and $ln\hat{Y}_T(1)$, so we use the forecast for period 78 for the forecast for period 79. And under expectation $\hat{\epsilon}_{T+2}$ becomes 0.

In this example, because $T = 77$, a Two-Step Ahead forecast is the forecast for period 79:

$$ln\hat{Y}_{77}(2) = 0.542ln\hat{Y}_{77}(1) + 0.458lnY_{77} + lnY_{67} - 0.542lnY_{66} - 0.458lnY_{65} - 0.795\hat{\epsilon}_{67}$$

We substitute in the appropriate values from Table 13.15 and the forecast for period 78 to determine a forecast for period 79.

$$ln\hat{Y}_{77}(2) = 0.542(5.633) + 0.458(5.606) + 5.781 - 0.542(5.549) - 0.458(5.242) - 0.795(-0.079)$$

$$ln\hat{Y}_{77}(2) = 6.056$$

And so on.

Three-Step Ahead Forecast for Period 80:

$$ln\hat{Y}_T(3) = 0.542ln\hat{Y}_T(2) + 0.458ln\hat{Y}_T(1) + lnY_{T-9} - 0.542lnY_{T-10} - 0.458lnY_{T-11} - 0.795\hat{\epsilon}_{T-9}$$

Four-Step Ahead Forecast for Period 81:

$$ln\hat{Y}_T(4) = 0.542ln\hat{Y}_T(3) + 0.458ln\hat{Y}_T(2) + lnY_{T-8} - 0.542lnY_{T-9} - 0.458lnY_{T-10} - 0.795\hat{\epsilon}_{T-8}$$

$\vdots$

Table 13.20

Obs		Natural Log Data	Fitted Value	Residual	Forecast
		⋮	⋮	⋮	
64		5.775	5.398	0.377	
65		5.242	5.457	-0.216	
66		5.549	5.557	-0.008	
67	JUL70	5.781	5.860	-0.079	
68		6.001	6.166	-0.165	
69		6.518	6.433	0.085	
70		6.755	6.756	-0.001	
71		6.797	6.846	-0.049	

72		6.498	6.510	-0.012	
73	JAN71	6.443	6.516	-0.074	
74		5.730	6.034	-0.304	
75		5.781	5.614	0.167	
76		5.513	5.457	0.056	
77		5.606	5.350	0.256	
78					5.633
79	JUL71				6.056
80					6.320
81					6.749
82					7.027
83					7.089
84					6.791
85	JAN72				

Figure 13-6

Figure 13-7

Returning to the Original Sales Values

We have forecasted the log of the Sales but we are really interested in forecasts in the original Sales values. Consequently, we take the exponential values of the Fitted Values and of the Forecasts.

For example, for Period 64, the Fitted Value for Log of Sales is 5.775, thus

$$exp(5.775) = e^{5.775} = 220.9$$

The Actual Value for Period 64 is 322, so the Residual is:

$$\text{Residual} = 322 - 220.9 = 101.1$$

Similary, the first forecast for Period 78 was a forecast of the Log of Sales:

$$\text{Forecast for Period 78} = 5.633\text{, hence}$$
$$\text{Sales Forecast for Period 78} = exp(5.633) = e^{5.633} = 279.5$$

And so on.

Table 13.21

Obs		Data	Exponential of Fitted Values	Residuals	Exponential of Forecasts
		⋮	⋮	⋮	
64		322	220.9	101.1	
65		189	234.5	-45.5	
66		257	259.0	-2.0	
67	JAN70	324	350.6	-26.6	
68		404	476.4	-72.4	
69		677	622.1	54.9	
70		858	858.8	-0.8	
71		895	940.1	-45.1	
72		664	672.1	-8.1	
73	JAN71	628	676.2	-48.2	
74		308	417.4	-109.4	
75		324	274.1	49.9	
76		248	234.4	13.6	

77		272	210.7	61.3	
78					279.5
79	JUL71				426.5
80					555.6
81					853.2
82					1,126.2
83					1,199.1
84					889.9
85	JAN72				

Figure 13-8

Figure 13-9

Confidence Intervals

A confidence interval is based on forecast variance. In the setting of *ARIMA* models the primary source of forecast variance is from the error term.

A Beginning Example

As an example, let us consider an *ARIMA*(0,1,1) model.

An *ARIMA*(0,1,1) model written in backshift notation is:

$$(1 - B)Y_T = (1 - \theta_1 B)\epsilon_T$$

In expanded notation this is:

$$Y_T = Y_{T-1} - \theta_1\epsilon_{T-1} + \epsilon_T$$

One-step ahead forecast confidence interval

A one-step ahead forecast, $\hat{Y}_T(1)$, is then the expectation of Y_{T+1}. Using the above formula and taking expectation we have

Actual $$Y_{T+1} = Y_T - \theta_1\epsilon_T + \epsilon_{T+1}$$

$$E(\mathrm{Y}_{T+1}) = E(\mathrm{Y}_T - \theta_1\epsilon_T + \epsilon_{T+1})$$

Forecast $$\hat{Y}_T(1) = \mathrm{Y}_T - \theta_1\epsilon_T$$

The Forecast Error is then just the difference between the Actual and the Forecast.

$$Error = Actual - Forecast$$

$$e_T(1) = \mathrm{Y}_{T+1} - \hat{Y}_T(1)$$

$$\mathrm{e}_T(1) = (\mathrm{Y}_T - \theta_1\epsilon_T + \epsilon_{T+1}) - (\mathrm{Y}_T - \theta_1\epsilon_T)$$

$$e_T(1) = \epsilon_{T+1}$$

The mean of the forecast error is zero, because by expectation

$$E\Big(e_T(1)\Big) = \mathrm{E}(\epsilon_{T+1})$$

$$E\Big(e_T(1)\Big) \;=\; 0$$

The variance of forecast error is thus the expectation of the squared forecast error.

$$VAR\Big(e_T(1)\Big) \;=\; \mathrm{E}(\epsilon^2_{T+1})$$

$$VAR\Big(e_T(1)\Big) \;=\; \sigma^2_\epsilon$$

One-step ahead forecast variance

$\boldsymbol{\sigma^2_{\hat{Y}_T(1)} \;=\; \sigma^2_\epsilon}$	**13.22**

Thus, the ***standard error of forecast*** is

$$\sqrt{\sigma^2_{\hat{Y}_T(1)}} \;=\; \sqrt{\sigma^2_\epsilon}$$

One-step ahead standard error of forecast

$\boldsymbol{\sigma_{\hat{Y}_T(1)} = \sigma_\epsilon}$	**13.23**

Indeed, for every one-step ahead forecast, the standard error of forecast is $\boldsymbol{\sigma_\epsilon}$.

Thus, the $(1 \;-\; \alpha)$% confidence interval around the one-step ahead forecast is

$$\hat{Y}_T(1) \;\pm\; \mathrm{t}_{\frac{\alpha}{2}}\,\sigma_{\hat{Y}_T(1)}$$

$$\hat{Y}_T(1) \;\pm\; \mathrm{t}_{\frac{\alpha}{2}}\sigma_\epsilon \qquad 13.24$$

Because the most common confidence interval is a 95% confidence interval, with a sample size larger than 30 it is reasonable to use 1.96 for $t_{\frac{\alpha}{2}}$, thus equation 13.24 reduces to:

One-step ahead forecast confidence interval

$\boldsymbol{\hat{Y}_T(1) \;\pm\; 1.96\sigma_\epsilon}$	**13.25**

Returning to the data and model used in the $ARIMA(0,1,1)$ example. The original data have $T = 30$ observations. Through differencing they are reduced to $T - 1 = 29$ observations, and the model:

$$\hat{Y}_t(1) \;=\; Y_t \;-\; .9184\hat{\epsilon}_{tt}$$

Table 13.22

Period t	Actual Y_t	Fit $\hat{Y}_t$	Residual $\hat{\epsilon}_t$	Squared Residual $\hat{\epsilon}^2_t$
1	354	355.6		
2	368	355.5	12.5	156.25
3	329	356.5	-27.5	756.25

4	389	354.2	34.8	1,211.04
⋮	⋮	⋮	⋮	⋮
24	345	353.4	-8.4	70.56
25	367	352.7	14.3	204.49
26	380	353.9	26.1	681.21
27	387	356.0	31.0	961.00
28	346	358.5	-12.5	156.25
29	321	357.5	-36.5	1,332.25
30	372	354.5	17.5	306.25

$$SSE = \sum \hat{\epsilon}^2 = 12{,}793.8$$

$$MSE = \frac{SSE}{29-1} = \frac{12{,}793.8}{28} = 456.9$$

$$\sigma_\epsilon = \sqrt{MSE} = \sqrt{456.9} = 21.38$$

$$\hat{Y}_T(1) \pm 1.96\sigma_\epsilon$$

$$355.9 \pm 1.96(21.38)$$
$$355.9 \pm 41.90$$
$$(\ 314.00 \qquad 397.80)$$

Two-step ahead forecast confidence interval

For a Two-Step Ahead Confidence Interval we must compare the Two-Step Ahead Forecast with the Two-Step Ahead Actual. The Two-Step Ahead Forecast, $\hat{Y}_T(2)$, is determined by taking expectation of the Two-Step Ahead Actual, Y_{T+2}

Actual $\qquad Y_{T+2} = Y_{T+1} - \theta_1\epsilon_{T+1} + \epsilon_{T+2}$

Forecast $\qquad E(\mathrm{Y}_{T+2}) = E(\mathrm{Y}_{T+1} - \theta_1\epsilon_{T+1} + \epsilon_{T+2})$

$$\hat{Y}_T(2) = \hat{Y}_T(1)$$

$$\hat{Y}_T(2) = Y_T - \theta_1\epsilon_T$$

Again, forecast error is the difference between the Actual value and the Forecast value.

$$Error = Actual - Forecast$$

$$e_T(2) = Y_{T+2} - \hat{Y}_T(2)$$

$$\mathrm{e}_T(2) = (Y_{T+1} - \theta_1\epsilon_{T+1} + \epsilon_{T+2}) - (Y_T - \theta_1\epsilon_T)$$

Because $Y_{T+1} = Y_T - \theta_1\epsilon_T + \epsilon_{T+1}$ by substitution we have

$$e_T(2) = [(Y_T - \theta_1\epsilon_T + \epsilon_{T+1}) - (\theta_1\epsilon_{T+1} + \epsilon_{T+2})] - (Y_T - \theta_1\epsilon_T)$$

$$e_T(2) = (1 - \theta_1)\epsilon_{T+1} + \epsilon_{T+2} \qquad 13.26$$

Again the mean of the forecast error is zero, so the forecast variance is thus the expectation of the squared forecast error. Or, the expectation of the square of equation 13.27.

$$\sigma^2_{\hat{Y}_T(2)} = E[e_T(2)]^2$$

$$= E[(1-\theta_1)\epsilon_{T+1} + \epsilon_{T+2}]^2$$

$$= E[(1-\theta_1)^2\epsilon^2_{T+1} + 2(1-\theta_1)\epsilon_{T+1}\epsilon_{T+2} + \epsilon^2_{T+2}]$$

$$= (1-\theta_1)^2\sigma^2_\epsilon + \sigma^2_\epsilon$$

Two-step ahead forecast variance for an *ARIMA*(0,1,1) model

$$\boldsymbol{\sigma^2_{\hat{Y}_T(2)} = \sigma^2_\epsilon\,[1 + (1-\theta_1)^2]} \qquad \textbf{13.27}$$

Then the standard error of forecast becomes

$$\sigma_{\hat{Y}_T(2)} = \sqrt{\sigma^2_\epsilon\,[1 + (1-\theta_1)^2]}$$

Or,

$$\sigma_{\hat{Y}_T(2)} = \sigma_\epsilon\sqrt{1 + (1-\theta_1)^2}$$

And thus a confidence interval for a two-step ahead forecast is

$$\hat{Y}_T(2) \pm t_{\frac{\alpha}{2}}\sigma_{\hat{Y}_T(2)}$$

$$\hat{Y}_T(2) \pm t_{\frac{\alpha}{2}}\sigma_\epsilon\sqrt{1 + (1-\theta_1)^2} \qquad 13.28$$

And in particular a 95% confidence interval is

Two-step ahead forecast confidence interval for an *ARIMA*(0,1,1) model

$$\boldsymbol{\hat{Y}_T(2) \pm 1.96\sigma_\epsilon\sqrt{1 + (1-\theta_1)^2}} \qquad \textbf{13.29}$$

Again, use the example data, we have

$$\hat{\theta}_1 = .9184$$

in $\hat{Y}_t(1) = Y_t - .9184\hat{\epsilon}_{tt}$

and $\sigma_\epsilon = 21.38$

$$\hat{Y}_T(2) \pm 1.96\sigma_\epsilon\sqrt{1 + (1-\theta_1)^2}$$

$$355.9 \pm 1.96(21.38)\sqrt{1 + (1-.9184)^2}$$

$$355.9 \pm 42.04$$

$$(\,313.86 \qquad 397.94)$$

Multi-step forecast confidence intervals

In similar fashion it can be shown that the three-step ahead forecast error variance is

$$\sigma^2_{\hat{Y}_T(3)} = \sigma^2_\epsilon[1 + 2(1-\theta_1)^2]$$

And thus a three-step ahead 95% confidence interval is

$$\hat{Y}_T(3) \pm 1.96\sigma_\epsilon\sqrt{1 + 2(1-\theta_1)^2} \quad 13.30$$

Continuing, a four-step ahead variance of forecast is

$$\sigma^2_{\hat{Y}_T(4)} = \sigma^2_\epsilon[1 + 3(1-\theta)^2]$$

The 95% confidence interval of forecast is

$$\hat{Y}_T(4) \pm 1.96\sigma_\epsilon\sqrt{1 + 3(1-\theta_1)^2} \quad 13.31$$

By now the pattern should be apparent. Forecast variances are:

$$\sigma^2_{\hat{Y}_T(1)} = \sigma^2_\epsilon$$

$$\sigma^2_{\hat{Y}_T(2)} = \sigma^2_\epsilon[1 + (1-\theta_1)^2]$$

$$\sigma^2_{\hat{Y}_T(3)} = \sigma^2_\epsilon[1 + 2(1-\theta_1)^2]$$

$$\sigma^2_{\hat{Y}_T(4)} = \sigma^2_\epsilon[1 + 3(1-\theta_1)^2]$$

$$\vdots$$

$$\sigma^2_{\hat{Y}_T(\ell)} = \sigma^2_\epsilon[1 + (\ell - 1)(1-\theta_1)^2] \quad 13.32$$

And the corresponding 95% confidence intervals of forecast

$$\hat{Y}_T(1) \pm 1.96\sigma_\epsilon$$

$$\hat{Y}_T(2) \pm 1.96\sigma_\epsilon\sqrt{1 + (1-\theta_1)^2}$$

$$\hat{Y}_T(3) \pm 1.96\sigma_\epsilon\sqrt{1 + 2(1-\theta_1)^2}$$

$$\hat{Y}_T(4) \pm 1.96\sigma_\epsilon\sqrt{1 + 3(1-\theta_1)^2}$$

$$\vdots$$

95% Confidence Intervals for this *ARIMA*(0,1,1) model.

1-step ahead	$355.9 \pm 1.96(21.38)$
2-steps ahead	$355.9 \pm 1.96(21.38)\sqrt{1 + (1-.9184)^2}$
3-steps ahead	$355.9 \pm 1.96(21.38)\sqrt{1 + 2(1-.9184)^2}$
4-steps ahead	$355.9 \pm 1.96(21.38)\sqrt{1 + 3(1-.9184)^2}$

$$\vdots$$

Table 13.23

Period t	Actual Y_t	Fit $\hat{Y}_t$	Forecast $\hat{Y}_T(\ell)$	95% Confidence Limits Lower	Upper
$\vdots$	$\vdots$	$\vdots$			
28	346	358.5			
29	321	357.5			
30	372	354.5			
31			355.9	314.0	397.9
32			355.9	313.9	398.0
33			355.9	313.8	398.1
34			355.9	313.6	398.3
35			355.9	313.5	398.4
36			355.9	313.4	398.5

Figure 13-10

ℓ step ahead forecast confidence interval for an *ARIMA*(0,1,1) model

$$\hat{Y}_T(\ell) \pm 1.96\sigma_\epsilon\sqrt{1 + (\ell - 1)(1-\theta_1)^2} \qquad \textbf{13.33}$$

Recall, though, that these confidence intervals were developed specifically for an *ARIMA*(0,1,1) model. Different models will have different confidence interval formulations.

We have noted before that an *ARIMA*(0,1,1) is equivalent to single exponential smoothing. Notice too then that if we set the smoothing constant α to $1 - \theta_1$, the above series of confidence intervals can be re-written.

$$\hat{Y}_T(1) \pm 1.96\sigma_\epsilon$$

$$\hat{Y}_T(2) \pm 1.96\sigma_\epsilon\sqrt{1 + \alpha^2}$$

$$\hat{Y}_T(3) \pm 1.96\sigma_\epsilon\sqrt{1 + 2\alpha^2}$$

$$\hat{Y}_T(4) \pm 1.96\sigma_\epsilon\sqrt{1 + 3\alpha^2}$$

$$\vdots$$

This series of equations is identical with the series of equations in Chapter 8, pages .

A Second Example of Determine Forecast Confidence Intervals

In this example we shall use the Sales Volume data (see page 11). We have $T = 48$ observations, so a one-step ahead forecast is a forecast for period 49,

One-step ahead forecast error

The one-step ahead actual is:

$$Y_{49} = Y_{48} + Y_{37} - Y_{36} - .556\hat{\epsilon}_{48} - .674\hat{\epsilon}_{37} + .375\hat{\epsilon}_{36} + \hat{\epsilon}_{49}$$

The one-step ahead forecast is:

$$\hat{Y}_{48}(1) = \mathrm{Y}_{48} + \mathrm{Y}_{37} - \mathrm{Y}_{36} - .556\hat{\epsilon}_{48} - .674\hat{\epsilon}_{37} + .375\hat{\epsilon}_{36}$$

Thus the one-step ahead forecast error is:

$$e_{48}(1) = Y_{49} - \hat{Y}_{48}(1)$$

$$\mathrm{e}_{48}(1) = \hat{\epsilon}_{49}$$

One-step ahead forecast variance

The variance of the one-step ahead forecast error is:

$$\sigma^2_{\hat{Y}_{48}(1)} = \sigma^2_{\epsilon}$$

From the computer output we have

$$\sigma^2_{\epsilon} = MSE = 659.1$$

We know always that the one-step ahead forecast variance is the *MSE* of the model.

$$\sigma^2_{\hat{Y}_{48}(1)} = \sigma^2_{\epsilon} = 659.1$$

Thus the standard error of the one-step ahead forecast is:

$$\sigma_{\hat{Y}_{48}(1)} = \sqrt{659.1} = 25.67$$

One-step ahead forecast confidence interval

$$\hat{Y}_T(1) \pm 1.96\sigma_{\epsilon}$$

$$592.96 \pm 1.96(25.67)$$

$$592.96 \pm 50.32$$

Table 13.24

Period	*Actual*	*Forecast*	*95% Confidence Limits*	
⋮	⋮		*Lower*	*Upper*
46	590.0			
47	636.4			
48	677.4			
49		592.96	542.64	643.28

Two-step ahead forecast error

The two-step ahead actual is:

$$Y_{50} = Y_{49} + Y_{38} - Y_{37} - .556\hat{\epsilon}_{49} - .674\hat{\epsilon}_{38} + .375\hat{\epsilon}_{37} + \hat{\epsilon}_{50}$$

The two-step ahead forecast is:

$$\hat{Y}_{48}(2) = \hat{\mathrm{Y}}_{48}(1) + \mathrm{Y}_{38} - \mathrm{Y}_{37} - .674\hat{\epsilon}_{38} + .375\hat{\epsilon}_{37}$$

Thus the two-step ahead forecast error is:

$$e_{48}(2) = Y_{50} - \hat{Y}_{48}(2)$$

$$e_{48}(2) = \left(Y_{49} - \hat{Y}_{48}(1)\right) - .556\hat{\epsilon}_{49} + \hat{\epsilon}_{50}$$

$Y_{49} - \hat{Y}_{48}(1)$ is the one-step forecast error

$$e_{48}(1) = \hat{\epsilon}_{49}$$

$$e_{48}(2) = \hat{\epsilon}_{49} - .556\hat{\epsilon}_{49} + \hat{\epsilon}_{50}$$

$$e_{48}(2) = (1 - .556)\hat{\epsilon}_{49} + \hat{\epsilon}_{50} = .444\hat{\epsilon}_{49} + \hat{\epsilon}_{50}$$

Two-step ahead forecast variance

The variance of the two-step ahead forecast error is:

$$\sigma^2_{\hat{Y}_{48}(2)} = \sigma^2_{\epsilon}[1 + (.444)^2]$$

$$\sigma^2_{\hat{Y}_{48}(2)} = 659.1[1.197\,] = 769.03$$

Thus the standard error of the two-step ahead forecast is:

$$\sigma_{\hat{Y}_{48}(2)} = \sqrt{769.03} = 28.09$$

Two-step ahead forecast confidence interval

The 95% confidence interval is:

$$\hat{Y}_{48}(2) \pm 1.96\sigma_{\hat{Y}_{48}(2)}$$

$$612.71 \pm 1.96(28.09)$$

(557.65 667.77)

Table 13.25

Period	*Actual*	*Forecast*	*95% Confidence Limits*	
⋮	⋮		*Lower*	*Upper*
46	590.0			
47	636.4			
48	677.4			
49		592.96	542.64	643.28
50		612.71	557.65	667.77

Three-step ahead forecast error

The three-step ahead actual is:

$$Y_{51} = Y_{50} + Y_{39} - Y_{38} - .556\hat{\epsilon}_{50} - .674\hat{\epsilon}_{39} + .375\hat{\epsilon}_{38} + \hat{\epsilon}_{51}$$

The three-step ahead forecast is:

$$\hat{Y}_{48}(3) = \hat{Y}_{48}(2) + Y_{39} - Y_{38} - .674\hat{\epsilon}_{39} + .375\hat{\epsilon}_{38}$$

Thus the three-step ahead forecast error is:

$$e_{48}(3) = Y_{51} - \hat{Y}_{48}(3)$$

$$e_{48}(3) \;=\; \left(Y_{50} \;-\; \hat{Y}_{48}(2)\right) \;-\; .556\hat{\epsilon}_{50} \;+\; \hat{\epsilon}_{51}$$

$Y_{50} \;-\; \hat{Y}_{48}(2)$ is the two-step ahead forecast error

$$e_{48}(2) \;=\; .444\hat{\epsilon}_{49} \;+\; \hat{\epsilon}_{50}$$

$$e_{48}(3) \;=\; \left(.444\hat{\epsilon}_{49} \;+\; \hat{\epsilon}_{50}\right) \;-\; .556\hat{\epsilon}_{50} \;+\; \hat{\epsilon}_{51}$$

$$e_{48}(3) \;=\; .444\hat{\epsilon}_{49} \;+\; .444\hat{\epsilon}_{50} \;+\; \hat{\epsilon}_{51}$$

Three-step ahead forecast variance

The three-step ahead forecast variance is:

$$\sigma^2_{\hat{Y}_{48}(3)} \;=\; \sigma^2_{\epsilon}[1 \;+\; (.444)^2 \;\;+\; (.444)^2]$$

$$\sigma^2_{\hat{Y}_{48}(3)} \;=\; 659.1[1.394\,] \;=\; 918.79$$

Thus the standard error of the three-step ahead forecast is:

$$\sigma_{\hat{Y}_{48}(3)} \;=\; \sqrt{918.79} \;=\; 30.31$$

Three-step ahead forecast confidence interval

The 95% confidence interval is:

$$\hat{Y}_{48}(3) \;\pm\; 1.96\sigma_{\hat{Y}_{48}(3)}$$

$$621.32 \;\pm\; 1.96(30.31)$$

$$(561.91 \qquad 680.73)$$

And so on. We determine all six forecast confidence intervals for the six forecasts.

⋮

Table 13.26

Period	*Actual*	*Forecast*	*95% Confidence Limits*	
⋮	⋮		*Lower*	*Upper*
46	590.0			
47	636.4			
48	677.4			
49		592.96	542.64	643.28
50		612.71	557.65	667.77
51		621.32	561.91	680.73
52		637.00	573.51	700.49
53		611.64	544.33	678.95
54		651.48	580.56	722.41

Figure 13-11

Figure 13-12

ψ-Weights

A somewhat more general approach to confidence intervals is through the use of ψ-*weights*. Roughly speaking the ψ-weights are the coefficients of a purely *MA* model. For example, let us consider an *MA*(2) model with no mean.

ARIMA(0,0,2)

$$Y_T = \epsilon_T - \theta_1\epsilon_{T-1} - \theta_2\epsilon_{T-2}$$

In ψ-weight form this equation is written:

$$Y_T = \psi_0\epsilon_T + \psi_1\epsilon_{T-1} + \psi_2\epsilon_{T-2}$$

The ψ-weights of this model are

$$\psi_0 = 1$$
$$\psi_1 = -\theta_1$$
$$\psi_2 = -\theta_2$$

Returning to the *ARIMA*(0,1,1) model we re-write it as a purely *MA* model.

$$Y_T = Y_{T-1} - \theta_1\epsilon_{T-1} + \epsilon_T \qquad 13.34$$

Shifting back equation 13.34 one-step we get

$$Y_{T-1} = Y_{T-2} - \theta_1\epsilon_{T-2} + \epsilon_{T-1} \qquad 13.35$$

Substituting equation 13.35 into equation 13.34

$$Y_T = [Y_{T-2} - \theta_1\epsilon_{T-2} + \epsilon_{T-1}] - \theta_1\epsilon_{T-1} + \epsilon_T$$

$$Y_T = Y_{T-2} - \theta_1\epsilon_{T-2} + (1 - \theta_1)\epsilon_{T-1} + \epsilon_T \qquad 13.36$$

In similar fashion Y_{T-2} is

$$Y_{T-2} = Y_{T-3} - \theta_1\epsilon_{T-3} + \epsilon_{T-2}$$

Then by substitution in equation 13.36

$$Y_T = [Y_{T-3} - \theta_1\epsilon_{T-3} + \epsilon_{T-2}] - \theta_1\epsilon_{T-2} + (1 - \theta_1)\epsilon_{T-1} + \epsilon_T$$

$$Y_T = Y_{T-3} - \theta_1\epsilon_{T-3} + (1 - \theta_1)\epsilon_{T-2} + (1 - \theta_1)\epsilon_{T-1} + \epsilon_T \qquad 13.37$$

Re-arranging terms of equation 13.37

$$Y_T = \epsilon_T + (1 - \theta_1)\epsilon_{T-1} + (1 - \theta_1)\epsilon_{T-2} - \theta_1\epsilon_{T-3} + Y_{T-3} \qquad 13.38$$

The pattern becomes clear, that if we continue to substitute values for Y_{T-h}

$$Y_T = \epsilon_T + (1 - \theta_1)\epsilon_{T-1} + (1 - \theta_1)\epsilon_{T-2} + (1 - \theta_1)\epsilon_{T-3} + \dots + (1 - \theta_1)\epsilon_{T-h} + \dots \qquad 13.39$$

In this particular case of an *ARIMA*(0,1,1), the ψ-weights are identical after ψ_0:

$$\psi_0 = 1$$

$$\psi_1 = 1 - \theta_1$$

$$\psi_2 = 1 - \theta_1$$

$$\vdots$$

$$\psi_h = 1 - \theta_1$$

$$\vdots$$

Forecast Confidence Intervals using ψ-Weights

It is not difficult to show that the variance of forecast error can be written in terms of ψ-weights.

$$\sigma^2_{\hat{Y}_T(1)} = \sigma^2_\epsilon$$

$$\sigma^2_{\hat{Y}_T(2)} = \sigma^2_\epsilon[1 + \psi_1^2]$$

$$\sigma^2_{\hat{Y}_T(3)} = \sigma^2_\epsilon[1 + \psi_1^2 + \psi_2^2]$$

$$\sigma^2_{\hat{Y}_T(4)} = \sigma^2_\epsilon[1 + \psi_1^2 + \psi_2^2 + \psi_3^2]$$

$$\vdots$$

$$\boldsymbol{\sigma^2_{\hat{Y}_T(\ell)} = \sigma^2_\epsilon[1 + \psi_1^2 + \psi_2^2 + \psi_3^2 + \ldots + \psi_{\ell-1}^2]} \qquad \textbf{13.40}$$

Equation 13.40 is completely general. Once the ψ-weights have been determined for a model, then equation 13.40 will produce the variance of forecast, and the correspondingly the confidence intervals.

ψ-weights can be determined using simple polynomial division. We use the backshift notation to determine the ψ-weights. Returning to our example of an *ARIMA*(0,1,1). In backshift notation;

$$(1 - \mathrm{B})\mathrm{Y}_T = (1 - \theta_1\mathrm{B})\epsilon_T$$

Algebraically, this can be re-written as

$$Y_T = \frac{(1-\theta_1\mathrm{B})}{(1-\mathrm{B})}\epsilon_T \qquad 13.41$$

The fraction, $\frac{(1-\theta_1 B)}{(1-\mathrm{B})}$, can be considered as polynomial division

$$1 - B\,\overline{\big)\,1 - \theta_1 B - 0B^2 - 0B^3 - \cdots}$$

Using polynomial division of the backshift operator we have

$$\frac{(1 - \theta_1\mathrm{B})}{(1 - \mathrm{B})} = 1 + (1 - \theta_1)\mathrm{B} + (1 - \theta_1)\mathrm{B}^2 + (1 - \theta_1)\mathrm{B}^3 + \ldots \qquad 13.42$$

Substituting 13.42 into 13.41 we have

$$Y_T = [1 + (1 - \theta_1)\mathrm{B} + (1 - \theta_1)\mathrm{B}^2 + (1 - \theta_1)\mathrm{B}^3 + \ldots]\epsilon_T$$

$$Y_T = \epsilon_T + (1 - \theta_1)\epsilon_{T-1} + (1 - \theta_1)\epsilon_{T-2} + (1 - \theta_1)\epsilon_{T-3} + \ldots + (1 - \theta_1)\epsilon_{T-h} + \ldots$$

Look familiar?

PROBLEMS AND QUESTIONS

Step 5. Forecasting **Forecasting and Confidence Intervals**

13.1 The data, fitted values and residuals below are from an *ARIMA*(1, 0, 0) model in which $\hat{\delta} = 62$ and $\hat{\phi}_1 = 0.7$

Period	Y_t	$\hat{Y}_t$	$\hat{\epsilon}_t$	$\hat{Y}_t(\ell)$
⋮	⋮	⋮	⋮	
45	215	210.8	4.2	
46	212	216.7	-4.7	
47	210			
48	212			
49				
50				
51				
52				

(a) Write down the model in backshift notation.
(b) Write down the model in expanded notation.
(c) Determine the fitted values for periods 48 and 49.
(d) Determine forecasts for 1, 2, 3, and 4-steps ahead.

13.2 The data, fitted values and residuals below are from an *ARIMA*(0, 1, 1) model in which $\hat{\theta}_1 = -0.5$

Period	Y_t	$\hat{Y}_t$	$\hat{\epsilon}_t$	$\hat{Y}_t(\ell)$
⋮	⋮	⋮	⋮	
76	137	142.8	-5.8	
77	146	141.5	4.5	
78	150	146.5	3.5	
79	159			
80	157			
81				
82				
83				
84				

(a) Write down the model in backshift notation.
(b) Write down the model in expanded notation.
(c) Determine the fitted values for periods 79 and 80.
(d) Determine forecasts for 1, 2, 3, and 4-steps ahead.

13.3 The data, fitted values and residuals below are from an *ARIMA*(1, 1, 1) model in which $\hat{\delta} = 0$, $\hat{\phi}_1 = 0.7$, and $\hat{\theta}_1 = -0.5$

Period	Y_t	$\hat{Y}_t$	$\hat{\epsilon}_t$	$\hat{Y}_t(\ell)$
⋮	⋮	⋮	⋮	
123	447	451.3	-4.2	
124	442	445.7	-3.7	
125	440			
126	442			
127				
128				
129				
130				

(a) Write down the model in backshift notation.
(b) Write down the model in expanded notation.
(c) Determine the fitted values for periods 125 and 126.
(d) Determine forecasts for 1, 2, 3, and 4-steps ahead.

13.4 The data, fitted values and residuals below are from an *ARIMA*(2, 1, 0) model in which $\hat{\phi}_1 = -1.4$ and $\hat{\phi}_2 = 0.5$

Period	Y_t	$\hat{Y}_t$	$\hat{e}_t$	$\hat{Y}_t(\ell)$
⋮	⋮	⋮	⋮	
174	605.7	611.5	-5.8	
175	612.5	608.0	4.5	
176	618.1	621.6	3.5	
177	624.0			
178	631.4			
179				
180				
181				
182				

(a) Write down the model in backshift notation.
(b) Write down the model in expanded notation.
(c) Determine the fitted values for periods 177 and 178.
(d) Determine forecasts for 1, 2, 3, and 4-steps ahead.