TS module 16 ARIMA(0,1,1) process

(The attached PDF file has better formatting.)

An ARIMA(0,1,1) process (= an IMA(1,1) process) is the cumulative sum of an MA(1) process. Final exam problems on ARIMA(0,1,1) processes can be solved by first principles.

For an MA(1) process, Cryer and Chan use $\theta$ instead of $\theta_1$. For consistency of notation with moving average processes of higher order (MA(2), MA(3), and so forth), the discussion forum postings often use $\theta_1$.

Read the subsection of MA(1) processes in Cryer and Chan, chapter 9, pages 197-198; know equation 9.3.21 on page 197. See also the equation at the bottom on page 202: for the IMA(1,1) model, $\psi_j = 1 - \theta$ for $j \geq 1$.

An exam problem may give the parameters of the MA(1) process of first differences. We need the most recent residual of the first differences to forecast future values. The residual of the first differences is the residual of the time series values. The exam problem may give the forecasted and actual values for the current period.

The exam problem may give only one moving average parameter, and it may give actual and forecast values of the original time series for several periods. We back into the values of $\mu$ and $\theta_1$. Exam problems use simple algebra to derive parameters, not the regression analysis used for the student projects.

Exercise 16.1: ARIMA(0,1,1) model

An ARIMA(0,1,1) process has $\theta = 0.8$, $\mu = 1$, and $\sigma^2_\varepsilon = 4$ as parameters of the MA(1) time series that is the first differences of the ARIMA process. The forecast for the most recent period (Period T) was $\hat{y}_T = 22.2$ and the actual value was $Y_T = 22.7$.

A. What is the residual of the ARIMA(0,1,1) process for the most recent period?
B. What is the residual of the underlying MA(1) process for the most recent period?
C. What is the one period ahead forecast for the underlying MA(1) process?
D. What is the one period ahead forecast for the ARIMA(0,1,1) process?
E. What is the two periods ahead forecast for the underlying MA(1) process?
F. What is the two periods ahead forecast for the ARIMA(0,1,1) process?
G. What is the variance of the one period ahead forecast for the MA(1) process?
H. What is the variance of the one period ahead forecast for the ARIMA(0,1,1) process?
I. What is the variance of the two periods ahead forecast for the MA(1) process?
J. What is the variance of the two periods ahead forecast for the ARIMA(0,1,1) process?

*Part A:* The residual of the ARIMA(0,1,1) process for Period T is 22.7 – 22.2 = 0.5.

*Part B:* The residual of the underlying MA(1) process for Period T is also 0.5.

The residual for an ARIMA process is also the residual for the underlying ARMA process.

- ARIMA forecast = actual ARMA value in the previous period + ARMA forecast.
- Actual ARIMA value = actual ARMA value in the previous period + actual ARMA value.

- The ARMA residual = the actual ARMA value – the ARMA forecast.
- The ARIMA residual = the actual ARIMA value – the ARIMA forecast
    - = the actual ARMA value – the ARMA forecast = the ARMA residual

*Part C:* The one period ahead forecast for the underlying MA(1) process is

$$\epsilon_{t+1} + \mu - \theta_1 \times \epsilon_t = 1 - 0.8 \times 0.5 = \epsilon_{t+1} + 0.6.$$

The mean of an error term $\epsilon$ is zero, so the forecast is 0.6.

*Part D:* The one period ahead forecast for the ARIMA(0,1,1) process is the most recent ARIMA value + the underlying ARMA forecast = 22.7 + 0.6 = 23.3. *Take heed:* The one period ahead forecast begins with the most recent value, not the most recent forecast.

*Part E:* The two period ahead forecast for the underlying MA(1) process is $\mu$ = 1. The expected value of each future residual is zero, so the forecast is the mean.

*Part F:* The two period ahead forecast for the ARIMA(0,1,1) process is the one period ahead ARIMA forecast + the two periods ahead ARMA forecast = 23.3 + 1 = 24.3.

*Part E:* The *k* period ahead forecast (*k* > 1) for the underlying MA(1) process is $\mu$ = 1.

*Part F:* The *k* period ahead forecast (*k* > 1) for the ARIMA(0,1,1) process is

$Y_T + \sum$MA(1) forecasts = 22.7 + 0.6 + (k – 1) × 1 = 32.3 + k.

*Part G:* The variance of the one period ahead forecast for the MA(1) process is

$$\text{var}(\epsilon_{t+1} + \mu - \theta_1 \times \epsilon_t) = \text{var}(\epsilon_{t+1}) + \text{var}(\mu - \theta_1 \times \epsilon_t)$$
$$= \text{var}(\epsilon_{t+1} + 0.6) = \text{var}(\epsilon_{t+1}) + \text{var}(0.6) = \text{var}(\epsilon_{t+1}) = \sigma^2_\varepsilon = 4$$

*Part H:* The variance of the one period ahead forecast for the ARIMA(0,1,1) process is

$$\text{var}(\epsilon_{t+1} + Y_T + \mu - \theta_1 \times \epsilon_t) = \text{var}(\epsilon_{t+1}) + \text{var}(Y_T + \mu - \theta_1 \times \epsilon_t)$$
$$= \text{var}(\epsilon_{t+1} + 23.3) = \text{var}(\epsilon_{t+1}) + \text{var}(23.3) = \text{var}(\epsilon_{t+1}) = \sigma^2_\varepsilon = 4$$

*Part I:* The variance of the two periods ahead forecast for the MA(1) process is

$$\text{var}(\epsilon_{t+2} + \mu - \theta_1 \times \epsilon_{t+1}) = \text{var}(\epsilon_{t+2} - \theta_1 \times \epsilon_{t+1}) + \text{var}(\mu)$$
$$= \text{var}(\epsilon_{t+1}) + \text{var}(-\theta_1 \times \epsilon_{t+1}) + \text{var}(1.0)$$
$$= \text{var}(\epsilon_{t+1}) + \theta_1^2 \times \text{var}(\epsilon_{t+1}) + \text{var}(1.0) = (1 + \theta_1^2) \times \text{var}(\epsilon_{t+1}) = 1.64 \times \sigma^2_\varepsilon = 6.560$$

*Part J:* The variance of the two periods ahead forecast for the ARIMA(0,1,1) process is

$$\text{var}(\epsilon_{t+2} + \mu - \theta_1 \times \epsilon_{t+1} + Y_T + \mu - \theta_1 \times \epsilon_t)$$

$$= \quad \text{var}(\epsilon_{t+2} + (1-\theta_1) \times \epsilon_{t+1}) + \text{var}(Y_T + 2 \times \mu - \theta_1 \times \epsilon_t)$$

$$= \quad \text{var}(\epsilon_{t+2}) + \text{var}(1-\theta_1 \times \epsilon_{t+1}) + \text{var}(24.3)$$

$$= \quad \text{var}(\epsilon_{t+2}) + (1-\theta_1)^2 \times \text{var}(\epsilon_{t+1}) = \{ 1 + (1-\theta_1)^2\} \times \text{var}(\epsilon_t) = 1.04 \times \sigma^2_{\varepsilon} = 4.160$$

*Jacob:* How does the IMA(1,1) process differ from a random walk?

- For a white noise process, $\theta_1 = 0$ so $\theta_1^2 = 0$. The variance of each forecast is $\sigma^2_\varepsilon$.
- The integrated series is a random walk, for which the variance of the *k* periods ahead forecast is $k \times \sigma^2_\varepsilon$.

*Rachel:* The difference is the correlation of the time series observations.

- For a random walk, the variances of each period are uncorrelated.
- For an ARIMA(0,1,1) model, the variances of adjoining periods are correlated.

Suppose the residual for the one period ahead value of the first differences is +1.

- The residual for the one period ahead value of the original time series is also +1.
- The residual for the two period ahead value of the original time series is $+1 - 0.8 \times +1 = +0.2$.

Let y be the time series of first differences. The residual of +1 in period t+1 means that $y_{t+1}$ increases by +1. Since {y} is an MA(1) model, $y_{t+2}$ decreases by $+1 \times -\theta_1$. The two periods ahead value of the integrated time series increases by $1 + 1 \times -\theta_1$.

*Intuition:* Suppose $\theta_1 = 1$. Every random error is exactly offset the next period. If the residual in period *t* is +5, so the value in period *t* increases by 5, and the value in period *t+1* decreases by 5. For the integrated (original) time series, the residual of +5 in period *t* causes the value in period *t* to increase by +5 and no change in the value in period *t+1*.

We consider the variance of the two periods ahead forecast. For simplicity, suppose the MA(1) model of first differences has a mean of zero ($\mu = \theta_0 = 0$).

- The one period ahead value increases by $\epsilon_{t+1}$.
- The two periods ahead value increases by $\epsilon_{t+2}$ and decreases by $0.8 \times \epsilon_{t+1}$.

For the integrated (original) time series

- The one period ahead value increases by $\epsilon_{t+1}$.
- The two periods ahead value increases by $\epsilon_{t+2} + (1 - 0.8) \times \epsilon_{t+1}$.

We extend this reasoning to all future forecasts:

For the integrated (original) time series

- The one period ahead value increases by $\epsilon_{t+1}$.
- The two periods ahead value increases by $\epsilon_{t+2} + (1 - 0.8) \times \epsilon_{t+1}$.
- The three periods ahead value increases by $\epsilon_{t+3} + (1 - 0.8) \times \epsilon_{t+2} + (1 - 0.8) \times \epsilon_{t+1}$.

The MA(1) model is stationary, so all the error terms has the same variance.

- The variance of the two periods ahead forecast is $\sigma^2_\varepsilon + (1 - 0.8)^2 \times \sigma^2_\varepsilon$.
- The variance of the three periods ahead forecast is $\sigma^2_\varepsilon + 2 \times (1 - 0.8)^2 \times \sigma^2_\varepsilon$.
- The variance of the *k* periods ahead forecast is $\sigma^2_\varepsilon + (k - 1) \times (1 - 0.8)^2 \times \sigma^2_\varepsilon$.

ARMA and ARIMA residuals are the same.

- Suppose the one period ahead forecast for the ARMA process is 1 and the actual value is 2, giving a residual of 1.
- ⇛ The one period ahead forecast for the ARIMA process is $Y_t + 1$ and the actual value is $Y_t + 2$, giving a residual of 1.

The residual of 1 in the one period ahead forecast has two effects:

- The value of the first differences *increases* by 1 in the first forecast period.
- The value of the first differences *decreases* by 0.8 × 1 in the second forecast period.

The two effects on the original ARIMA time series are

- The value increases by 1 in the first forecast period.
- The value increases by 1 – 0.8 × 1 = 0.2 in the second forecast period.

For the ARIMA(0,1,1) process, each error term has a variance of $\sigma^2$. The effect on the forecast of lag *k* is

- $\sigma^2$ × the error term in future period *k*.
- $(1 - \theta_1)^2 \times \sigma^2$ × the error term in future period *k-1*.
- $(1 - \theta_1)^2 \times \sigma^2$ × the error term in future period *k-2*.
- …

The error terms are independent, so these variances are independent. The integrated ARIMA time series uses the sum of these variances. The variance of the *k* periods ahead forecast is $[1 + (k-1) \times (1 - \theta_1)^2] \times \sigma^2$.

*Take heed:* Remember the sign convention for the moving average parameters. A moving average parameter of $\theta_j$ adds $-\theta_j \times \epsilon_{t-j}$ in Period *t*.

- A *positive* $\theta_1$ in an ARIMA(0,1,1) model offsets the changes in adjacent periods.
  - For the *effect on the two periods ahead forecast*, a residual of 1 with a $\theta_1$ of 0.8 is like a residual of 0.2 and a $\theta_1$ of zero.
- A negative $\theta_1$ in an ARIMA(0,1,1) model causes parallel changes in adjacent periods.
  - For the *effect on the two periods ahead forecast*, a residual of 1 with a $\theta_1$ of –0.8 is like a residual of 1.8 and a $\theta_1$ of zero.

*Illustration:* ARIMA(0,1,1) with $\theta_1 = 1$:

If $\theta_1 = 1$, a residual in Period *t* causes an equal and opposite change in Period *t*+1. The only uncertainty in the ARIMA forecast is the residual in the forecast period.