TS Module 18 Forecast updates intuition

(The attached PDF file has better formatting.)

Exam problems relating ARIMA coefficients to forecasts may

- Give the ARIMA parameters and derive forecasts.
- Give forecasts and derive the ARIMA parameters.

Some exam problems do both. They give

- the forecasts for the next one or more periods
- the actual value in the next period
- ask for the new forecasts for the next one or more periods

To solve these problems

- Derive the ARIMA parameters from the old forecasts.
- Derive the new forecasts from the new values.

The exam problems generally give the type of process, such as AR(1) or ARMA(1,1).

- Some problems derive all the ARIMA parameters from the forecasts.
- Some problems give the mean and derive $\phi_1$ or $\theta_1$ (or other parameters).

*Take heed:* Suppose the exam problem gives the forecast for Period 101 and the actual observed value for Period 101. We use the observed value for the autoregressive part of the model and the residual for the moving average part of the model.

This posting uses the parameter $\theta_0$:

- The Cryer and Chan textbook uses $\theta_0$.
- The previous textbook for this course uses $\delta$.

Exercise 1.1: Revised Forecasts AR(1) Model

An AR(1) model of 300 observations has an estimated mean of 3, an estimated $\phi$ of 0.5, and $y_{300}$ = 2.6

- We *forecast* the next three periods: $\hat{y}_{300}(1)$, $\hat{y}_{300}(2)$, and $\hat{y}_{300}(3)$.
- In period 301, the *actual value* is 3.2.

A. What is estimated $\theta_0$ for this AR(1) model?
B. In Period 300, what are the forecasts for periods 301, 302, and 303, or $\hat{y}_{300}(1)$, $\hat{y}_{300}(2)$, and $\hat{y}_{300}(3)$?
C. If the actual value in Period 301 is 3.2, what are the revised forecasts for periods 302 and 303, or $\hat{y}_{301}(1)$ and $\hat{y}_{301}(2)$?

*Part A:* We use the estimated mean $\mu$ and autoregressive parameter $\phi$:

$$\theta_0 = \mu \times (1 - \varphi_1) = 3 \times (1 - 0.5) = 1.500$$

We use the same formula for all ARIMA models to derive $\theta_0$ from $\mu$ or $\mu$ from $\theta_0$.

*Part B:* We solve for the forecasts step by step:

- The forecast for Period 301 is 1.500 + 0.5 × 2.600 = 2.800.
- The forecast for Period 302 is 1.500 + 0.5 × 2.800 = 2.900.
- The forecast for Period 303 is 1.500 + 0.5 × 2.900 = 2.950.

We can also write the AR(1) process as a mean reversion, and use differences from the mean. An AR(1) model is $y_t = \theta_0 + \varphi_1 y_{t-1} + \varepsilon_t$. We write $\theta_0$ in terms of the mean $\mu$ as $\theta_0 = \mu \times (1 - \varphi_1)$. We rewrite the AR(1) formula as

$$y_t = \mu \times (1 - \varphi_1) + \varphi_1 y_{t-1} + \varepsilon_t \Rightarrow$$
$$y_t - \mu = \varphi_1 (y_{t-1} - \mu) + \varepsilon_t$$

- The difference from the mean shrinks each period by a constant proportion.
- The actual value is then distorted by a random fluctuation $\varepsilon_t$.

The shrinkage from the mean is called *mean reversion*.

- Mean reversion does *not* imply that the *actual values* get closer to the mean.
- Mean reversion implies that the *forecasts* get closer to the mean as the forecast interval increases.

The *forecasts* are scalars, not random variables; they have no error term. If $y_t - \mu = k$, then

- $\hat{y}_t(1) - \mu = k \times \varphi_1$

- $\hat{y}_t(2) - \mu = k \times \varphi_1^2$
- $\hat{y}_t(3) - \mu = k \times \varphi_1^3$
- $\hat{y}_t(n) - \mu = k \times \varphi_1^n$

We determine the forecasts in each future period:

- For Period 300, $y_t - \mu$, (the difference from mean) is 2.6 – 3.0 = –0.4
- For Period 301, $y_t - \mu$, (the difference from mean) is –0.4 × 0.5 = –0.2
- For Period 302, $y_t - \mu$, (the difference from mean) is –0.2 × 0.5 = –0.1
- For Period 303, $y_t - \mu$, (the difference from mean) is –0.1 × 0.5 = –0.05

*Part C:* In Period 301, the actual value is 3.2. We revise the forecasts for Periods 302 and 303.

- The forecast for Period 302 is 1.500 + 0.5 × 3.200 = 3.100.
- The forecast for Period 303 is 1.500 + 0.5 × 3.100 = 3.050.

Alternatively, we use differences from the mean. We start one period later, with the actual value in Period 301.

- For Period 301, $y_t - \mu$, (the difference from mean) is 3.2 – 3.0 = +0.2
- For Period 302, $y_t - \mu$, (the difference from mean) is +0.2 × 0.5 = +0.1
- For Period 303, $y_t - \mu$, (the difference from mean) is +0.1 × 0.5 = +0.05

The difference from the mean technique is often easier. On exam problems,

- subtract the mean from the given values
- derive the forecasts (and variances, standard deviations, and confidence intervals)
- add back the mean.

This avoids errors in adding $\theta_0$ to each term. If any autoregressive parameters are negative, using differences from the mean clarifies the pattern, since oscillations are around zero.

Exercise 1.2: AR(1) Forecasts

An AR(1) model of 90 day Treasury bill yields has a mean of 5.00%. The values are for the first day of each month.

- In December 20X7, we estimate the 90 day Treasury bill yield as 5.40% for January 20X8 and 5.05% for April 20X8.
- In January 20X7, the actual 90 day Treasury bill yield is 5.80%. We do not change the estimates for $\theta_0$ or $\varphi_1$.

A. What does it mean that an AR(1) model is mean reverting at a constant proportional rate?
B. What is $\varphi_1$ for this AR(1) model?
C. What are the original estimates (in December 20X7) for February 20X8 and March 20X8?
D. What are the revised estimates (in January 20X8) for February, March, and April 20X8?

*Part A:* An AR(1) model is $y_t = \theta_0 + \varphi_1 y_{t-1} + \varepsilon_t$.

We write $\theta_0$ in terms of the mean $\mu$ as $\theta_0 = \mu \times (1 - \varphi_1)$. We rewrite the AR(1) formula as

$$y_t = \mu \times (1 - \varphi_1) + \varphi_1 y_{t-1} + \varepsilon_t \Rrightarrow$$
$$y_t - \mu = \varphi_1 (y_{t-1} - \mu) + \varepsilon_t$$

The difference between the value and the mean shrinks each period by a constant proportion. The actual value is then distorted by a random fluctuation $\varepsilon_t$. The forecasts are scalars, not random variables; they have no error term. If $y_t - \mu = k$, then

- $\hat{y}_t(1) - \mu = k \times \varphi_1$
- $\hat{y}_t(2) - \mu = k \times \varphi_1^2$
- $\hat{y}_t(3) - \mu = k \times \varphi_1^3$
- $\hat{y}_t(n) - \mu = k \times \varphi_1^n$

Given the *current value and any forecast*, we derive $\varphi_1$ as $(\hat{y}_t(n) - \mu) / (y_t - \mu) = \varphi_1^n$

Given *any two forecasts*, we derive $\varphi_1$ as $(\hat{y}_t(n) - \mu) / ( \hat{y}_t(m) - \mu) = \varphi_1^{n-m}$

In this exercise,

- The one period ahead forecast is 5.4%, so the difference from the mean is 0.4%.
- The four periods ahead forecast is 5.05%, so the difference from the mean is 0.05%.

For the formula, n = 4 and m = 1, so we have

$0.05\% / 0.4\% = \varphi_1{}^3 \Rrightarrow \varphi_1 = [\, 0.05\% / 0.4\% \,]^{1/3} = 0.500$

*Take heed:* A third degree polynomial equation has three roots. This equation has one real root and two imaginary roots. We *discard the imaginary roots and choose the real root*. In general, if (n – m) is an odd number (1, 3, 5, …), we have one real root.

If (n – m) is an even number (2, 4, 6, …), we have two real roots. If $\varphi_1$ solves the equation, so does $-\varphi_1$.

The exercise may give the pattern of the forecasts or the pattern of the autocorrelations.

- If the forecasts approach the mean asymptotically, $\varphi_1$ is positive.
- If the forecasts oscillate about the mean, $\varphi_1$ is negative.

*Illustration:* In this exercise, had the problem given the forecasts for January and May, the difference 5 – 1 = 4 is even. We would not know the sign of $\varphi_1$. We would know the forecasts for March and July, but not the forecasts for February, April, and June.

*Part C:* The forecasted difference from the mean in January 20X8 is 5.40% – 5.00% = 0.40%. The forecasted differences from the mean for February and March 20X8 are

- February 20X8: 0.40% × 0.500 = 0.20%
- March 20X8: 0.20% × 0.500 = 0.10%

The forecasts add back the mean:

- February 20X8: 0.20% + 5.00% = 5.20%
- March 20X8: 0.10% + 5.00% = 5.10%

*Part D:* The actual difference from the mean in January 20X8 is 5.80% – 5.00% = 0.80%. The forecasted differences from the mean for February, March, and April 20X8 are

- February 20X8: 0.80% × 0.500 = 0.40%
- March 20X8: 0.40% × 0.500 = 0.20%
- April 20X8: 0.20% × 0.500 = 0.10%

The forecasts add back the mean:

- February 20X8: 0.40% + 5.00% = 5.40%
- March 20X8: 0.20% + 5.00% = 5.20%
- April 20X8: 0.10% + 5.00% = 5.10%

Exercise 1.3: AR(1) Forecasts

An AR(1) model of 90 day Treasury bill yields has a mean of 5.00%. The values are for the first day of each month.

- In December 20X7, we estimate the 90 day Treasury bill yield as 5.40% for January 20X8 and 5.05% for April 20X8.
- In January 20X7, the actual 90 day Treasury bill yield is 5.80%. We do not change the estimates for $\theta_0$ or $\varphi_1$.

A. What is $\varphi_1$ for this AR(1) model?
B. What are the original estimates (in December 20X7) for February 20X8 and March 20X8?
C. What are the revised estimates (in January 20X8) for February, March, and April 20X8?

*Part A:* An AR(1) model is $y_t = \theta_0 + \varphi_1\, y_{t-1} + \varepsilon_t$.

We write $\theta_0$ in terms of the mean $\mu$ as $\theta_0 = \mu \times (1 - \varphi_1)$. We rewrite the AR(1) formula as

$$y_t = \mu \times (1 - \varphi_1) + \varphi_1\, y_{t-1} + \varepsilon_t \Rrightarrow$$
$$y_t - \mu = \varphi_1\, (y_{t-1} - \mu) + \varepsilon_t$$

The difference between the value and the mean shrinks each period by a constant proportion. The actual value is then distorted by a random fluctuation $\varepsilon_t$. The forecasts are scalars, not random variables; they have no error term. If $y_t - \mu = k$, then

- $\hat{y}_t(1) - \mu = k \times \varphi_1$
- $\hat{y}_t(2) - \mu = k \times \varphi_1^{\,2}$
- $\hat{y}_t(3) - \mu = k \times \varphi_1^{\,3}$
- $\hat{y}_t(n) - \mu = k \times \varphi_1^{\,n}$

Given the *current value and any forecast*, we derive $\varphi_1$ as $(\hat{y}_t(n) - \mu) / (y_t - \mu) = \varphi_1^{\,n}$

Given *any two forecasts*, we derive $\varphi_1$ as $(\hat{y}_t(m) - \mu) / (\hat{y}_t(n) - \mu) = \varphi_1^{\,n-m}$

In this exercise,

- The one period ahead forecast is 5.4%, so the difference from the mean is 0.4%.
- The four periods ahead forecast is 5.05%, so the difference from the mean is 0.05%.

For the formula, n = 4 and m = 1, so we have

$0.05\% / 0.4\% = \varphi_1^{\,3} \Rrightarrow \varphi_1 = [\,0.05\% / 0.4\%\,]^{1/3} = 0.500$

*Take heed:* A third degree polynomial equation has three roots. This equation has one real root and two imaginary roots. We *discard the imaginary roots and choose the real root*. In general, if (n – m) is an odd number (1, 3, 5, …), we have one real root.

If (n – m) is an even number (2, 4, 6, …), we have two real roots. If $\varphi_1$ solves the equation, so does $-\varphi_1$.

The exercise may give the pattern of the forecasts or the pattern of the autocorrelations.

- If the forecasts approach the mean asymptotically, $\varphi_1$ is positive.
- If the forecasts oscillate about the mean, $\varphi_1$ is negative.

*Illustration:* In this exercise, had the problem given the forecasts for January and May, the difference 5 – 1 = 4 is even. We would not know the sign of $\varphi_1$. We would know the forecasts for March and July, but not the forecasts for February, April, and June.

*Part C:* The forecasted difference from the mean in January 20X8 is 5.40% – 5.00% = 0.40%. The forecasted differences from the mean for February and March 20X8 are

- February 20X8: 0.40% × 0.500 = 0.20%
- March 20X8: 0.20% × 0.500 = 0.10%

The forecasts add back the mean:

- February 20X8: 0.20% + 5.00% = 5.20%
- March 20X8: 0.10% + 5.00% = 5.10%

*Part D:* The actual difference from the mean in January 20X8 is 5.80% – 5.00% = 0.80%. The forecasted differences from the mean for February, March, and April 20X8 are

- February 20X8: 0.80% × 0.500 = 0.40%
- March 20X8: 0.40% × 0.500 = 0.20%
- April 20X8: 0.20% × 0.500 = 0.10%

The forecasts add back the mean:

- February 20X8: 0.40% + 5.00% = 5.40%
- March 20X8: 0.20% + 5.00% = 5.20%
- April 20X8: 0.10% + 5.00% = 5.10%

Question 1.4: ARIMA(0,1,1) process

An ARIMA(0,1,1) model for a time series of 25 observations, $y_t$, t = 1, 2, …, 25 has $\theta_1 = 0.6$, $\hat{y}_{24}(1) = 11$, $y_{25} = 12$, and $\hat{y}_{25}(1) = 13$.

A. What the mean $\mu$?
B. What is $\hat{y}_{25}(2)$?
C. If the actual observed value in Period 26 is 14, what is $\hat{y}_{26}(1)$?

*Part A:* The residual in Period 25 is 12 – 11 = 1. This is also the residual of the MA(1) process of the first differences. The forecast for the next period is $\mu$ (the mean of the first differences) – 0.6 × 1. The forecasted first difference is 13 – 12 = 1, so $\mu$ = 1 + 0.6 = 1.6.

*Part B:* In Period 25, the two periods ahead forecast of the first differences for Period 27 is the mean of 1.6. The observed value of the original time series in Period 26 is 13, so the forecast for Period 27 of the original time series is 13 + 1.6 = 14.6.

*Part C:* The observed value in Period 26 is 14, so the residual is 14 – 13 = 1. This is also the residual for the first differences. The forecasted first difference for Period 27 is 1.6 – 0.6 × 1 = 1. The observed value of the original time series in Period 26 is 14, so the forecasted value for Period 27 for the original time series is 14 + 1 = 15.