B.Sc./Grad. Dip.: Probability Models and Time Series
MAS programmes: Stochastic Models and Time Series

9 Forecasting for ARIMA models

9.1 Introduction

This chapter will further consider the theoretical aspects of making forecasts for data that can be modelled within the ARIMA paradigm. By way of introduction, we will first carry out forecasting for an AR(1) model (thus, in particular, this relates to the case of ARIMA(p,d,q) where $q = d = 0$ and $p = 1$); $d = 0$ implies that the model considered is stationary. We then consider the calibration and behaviour of the forecast function in the non-stationary case, in which $d \geq 1$. We close the chapter with an example of forecasting for data that is most appropriately modelled by a non-stationary ARIMA model.

9.2 Forecasting for the AR(1) model

Consider the AR(1) model

$$Y_t = \mu + \phi(Y_{t-1} - \mu) + \epsilon_t, \qquad t \in \mathbb{Z}, \tag{1}$$

and, to begin with, the forecast at lead time 1. Putting $t = T + 1$ in Equation (1),

$$Y_{T+1} = \mu + \phi(Y_T - \mu) + \epsilon_{T+1}.$$

Hence

$$\hat{y}_T(1) = E[Y_{T+1}|\mathcal{H}_T] = E[\mu + \phi(Y_T - \mu) + \epsilon_{T+1}|\mathcal{H}_T],$$

so that, since Y_T is determined by $\mathcal{H}_T$ and ϵ_{T+1} is independent of $\mathcal{H}_T$ with zero mean,

$$\hat{y}_T(1) = \mu + \phi(y_T - \mu). \tag{2}$$

The forecast error is given by

$$e_T(1) = Y_{T+1} - \hat{y}_T(1) = \epsilon_{T+1}. \tag{3}$$

Equation (3) expresses the fact that the white noise terms are the one step ahead forecast errors, a fact which is valid generally for this type of forecasting for ARIMA models. Hence the forecast error variance at lead time 1 is given by

$$V(1) = \text{var}(e_T(1)) = \text{var}(\epsilon_{T+1}) = \sigma^2, \tag{4}$$

where σ^2 is the variance of the underlying white noise process.

More generally, for any $h \geq 2$, putting $t = T + h$ in Equation (1), and taking conditional expectation, yields

$$\hat{y}_T(h) = E[Y_{T+h}|\mathcal{H}_T] = E[\mu + \phi(Y_{T+h-1} - \mu) + \epsilon_{T+h}|\mathcal{H}_T].$$

Noting that ϵ_{T+h} is independent of $\mathcal{H}_T$ with zero mean, we obtain the recurrence relation

$$\hat{y}_T(h) = \mu + \phi(\hat{y}_T(h-1) - \mu), \qquad h \geq 2. \tag{5}$$

If we define $\hat{y}_T(0) = y_T$, we can combine equations (2) and (5) to write

$$\hat{y}_T(h) = \mu + \phi(\hat{y}_T(h-1) - \mu), \qquad h \geq 1. \tag{6}$$

Equation (6) may be iterated numerically to obtain forecasts for successive lead times. Iterating Equation (6) algebraically, we obtain

$$\hat{y}_T(h) = \mu + \phi^h(y_T - \mu), \qquad h \geq 1. \tag{7}$$

Note that the influence of the observation y_T on the forecast $\hat{y}_T(h)$ declines geometrically as the lead time h increases.

Remarks 9.2.1
To derive Equation (7) by an alternative method and to obtain in addition an expression for the forecast error, use Equation (1) recursively with $t = T+h, T+h-1, \ldots, T+1$:

$$\begin{aligned} Y_{T+h} - \mu &= \phi(Y_{T+h-1} - \mu) + \epsilon_{T+h} \\ &= \phi^2(Y_{T+h-2} - \mu) + \phi\epsilon_{T+h-1} + \epsilon_{T+h} \\ \ldots \\ &= \phi^h(Y_T - \mu) + \sum_{i=0}^{h-1} \phi^i \epsilon_{T+h-i}. \end{aligned}$$

Thus

$$Y_{T+h} = \mu + \phi^h(Y_T - \mu) + \sum_{i=0}^{h-1} \phi^i \epsilon_{T+h-i}. \tag{8}$$

[More formally, we may give a proof of the result of Equation (8) by mathematical induction on h.]

Proposition 9.2.2 (Forecast error and variance for AR(1))
Consider the AR(1) model as specified by equation (1), and suppose that the origin and lead time for the forecasts are given by T and h, respectively. Then, the forecast error is given by

$$\sum_{i=0}^{h-1} \phi^i \epsilon_{T+h-i}$$

and the forecast error variance is given by

$$\frac{1-\phi^{2h}}{1-\phi^2}\sigma^2.$$

Proof
By putting $t = T+1$ in equation (1) we immediately obtain Equation (8) for the case $h = 1$.

Suppose now that the result of Equation (8) holds for some given $h \geq 1$. Putting $t = T+h+1$ in equation (1),

$$\begin{aligned}
Y_{T+h+1} &= \mu + \phi(Y_{T+h} - \mu) + \epsilon_{T+h+1} \\
&= \mu + \phi\left[\phi^h(Y_T - \mu) + \sum_{i=0}^{h-1} \phi^i \epsilon_{T+h-i}\right] + \epsilon_{T+h+1} \\
&= \mu + \phi^{h+1}(Y_T - \mu) + \sum_{i=0}^{h-1} \phi^{i+1} \epsilon_{T+h-i} + \epsilon_{T+h+1} \\
&= \mu + \phi^{h+1}(Y_T - \mu) + \sum_{i=0}^{h} \phi^i \epsilon_{T+h+1-i}.
\end{aligned}$$

Hence the result of Equation (8) holds also for $h+1$, which completes the proof by induction.

Taking expectations conditional upon $\mathcal{H}_T$ in Equation (8) and noting that the white noise terms $\epsilon_{T+h}, \epsilon_{T+h-1}, \ldots, \epsilon_{T+1}$ are independent of $\mathcal{H}_T$ with means zero, we re-derive Equation (7).

In addition, from Equations (7) and (8), we find an expression for the forecast error,

$$\begin{aligned}
e_T(h) &= Y_{T+h} - \hat{y}_T(h) \\
&= \sum_{i=0}^{h-1} \phi^i \epsilon_{T+h-i}. \qquad (9)
\end{aligned}$$

Thus the first two terms on the right hand side of Equation (8) represent the forecast and the final term is the forecast error. The corresponding forecast error variance is given by

$$\begin{aligned}
V(h) &= \mathrm{var}(e_T(h)) \\
&= \mathrm{var}\left(\sum_{i=0}^{h-1} \phi^i \epsilon_{T+h-i}\right) \\
&= \sum_{i=0}^{h-1} \phi^{2i} \sigma^2 \\
&= \frac{1-\phi^{2h}}{1-\phi^2}\,\sigma^2. \qquad (10)
\end{aligned}$$

□

Remarks 9.2.3 (Limiting behaviour of forecast)
Note that, as $h \to \infty$,

$$\hat{y}_T(h) \to \mu$$

and

$$V(h) \uparrow \frac{\sigma^2}{1-\phi^2} = \gamma_0.$$

9.3 Bread price example

In R, the function `predict` may be applied to an object of the class `Arima` to produce forecasts and their standard errors. Recall that the object `BP.ar1m` contains the results of fitting an AR(1) model, by the method of maximum likelihood, to the object `BP.rts`, which contains the bread price data. The function `predict` is applied to `BP.ar1m` with an additional parameter `10`, to indicate that forecasts for lead times up to 10 are to be produced. The ten forecasts produced, for the years 1758 to 1767, using 1757 as origin, are listed in `BP.fore$pred`. Finally, we calculate approximate 95% prediction limits at ±1.96 standard errors.

```
> BP.fore <- predict(BP.ar1m, 10)
> BP.fore
$pred
Time Series:
Start = 1758
End = 1767
Frequency = 1
 [1] 5.943184 5.842361 5.777539 5.735862 5.709067 5.691839 5.680762 5.673641
 [9] 5.669062 5.666119

$se
Time Series:
Start = 1758
End = 1767
Frequency = 1
 [1] 0.9303292 1.1060230 1.1709735 1.1967927 1.2073042 1.2116227 1.2134034
 [8] 1.2141386 1.2144425 1.2145680

> pred <- BP.fore$pred
> L95 <- pred - 1.96 * BP.fore$se
> U95 <- pred + 1.96 * BP.fore$se
> year<-1758:1767
> BP.PL95 <- data.frame(year,L95, pred, U95)
> BP.PL95
   year      L95     pred      U95
1  1758 4.119739 5.943184 7.766629
2  1759 3.674556 5.842361 8.010166
3  1760 3.482431 5.777539 8.072647
4  1761 3.390148 5.735862 8.081576
5  1762 3.342750 5.709067 8.075383
6  1763 3.317058 5.691839 8.066619
7  1764 3.302492 5.680762 8.059033
8  1765 3.293929 5.673641 8.053353
9  1766 3.288755 5.669062 8.049370
10 1767 3.285565 5.666119 8.046672
```

To check the above output, the average bread price in 1757 was $y_{124} = 6.1$. The estimates produced by R were $\hat{\phi} = 0.6429364$ and $\hat{\mu} = 5.6608179$ with $\hat{\sigma}^2 = 0.8655124$. Substituting these values into Equation (9) of Chapter 8,

$$\hat{y}_{124}(h) = 5.6608179 + (0.6429364)^h(6.1 - 5.6608179), \qquad h \geq 1.$$

In particular, the forecast average bread price for 1758 is

$$\hat{y}_{124}(1) = 5.6608179 + (0.6429364)(6.1 - 5.6608179) = 5.94318.$$

to 5 d.p.
From Equation (7) of Chapter 8, the forecast error variance is

$$V(1) = \hat{\sigma}^2 = 0.8655124.$$

Given the assumed normality of the data, the 95% prediction interval is

$$5.94318 \pm 1.96\sqrt{0.8655124} = 5.94318 \pm 1.8234452 = (4.120, 7.767).$$

The forecast and the prediction interval agree with those in the R output. The actual average price in 1758 was in fact 5.2.

→ Remember, apply differencing to Y_t and call that W_t, then do ARMA for W_t, then sub Y_t back in for ARIMA.

9.4 Forecasting for the ARIMA(0,1,1) model

9.4.1 Motivating discussion

$Y_t = \mu + \sum_{u=1} \phi_u (Y_{t-u} - \mu) + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$

Consider the ARIMA(0,1,1) model with mean zero,

Let $W_t = \Delta Y_t = (1-L)Y_t$

$W_t = \epsilon_t + \theta L \epsilon_t = (1+\theta L)\epsilon_t$

$$(1 - L)Y_t = (1 + \theta L)\epsilon_t, \tag{11}$$

$\Delta Y_t = (1-L)Y_t = (1+\theta L)\epsilon_t$

i.e., Model Eq ARIMA(0,1,1)

$$Y_t = Y_{t-1} + \epsilon_t + \theta\epsilon_{t-1}, \qquad t \in \mathbb{Z}. \tag{12}$$

Putting $t = T + h$ in Equation (12) and substituting into Equation (2) of Chapter 8,

$$\hat{y}_T(h) = E[Y_{T+h}|\mathcal{H}_T] = E[Y_{T+h-1} + \epsilon_{T+h} + \theta\epsilon_{T+h-1}|\mathcal{H}_T]. \tag{13}$$

$h=1$ Putting $h = 1$ in Equation (13) and noting that Y_T and ϵ_T are determined by $\mathcal{H}_T$ and ϵ_{T+1} is independent of $\mathcal{H}_T$ with mean zero, we obtain

$$\hat{y}_T(1) = y_T + \theta\epsilon_T.$$

$h \geq 2$ For $h \geq 2$, both ϵ_{T+h} and ϵ_{T+h-1} are independent of $\mathcal{H}_T$. Hence, from Equation (13), we obtain the recurrence relation

$$\hat{y}_T(h) = \hat{y}_T(h-1), \qquad h \geq 2.$$

Thus $\hat{y}_T(h)$ takes the same value for all lead times $h \geq 1$,

wow! the prediction for different future values (and differentially by) is the same!

what is ϵ_t?

$$\hat{y}_T(h) = y_T + \theta\epsilon_T, \qquad h \geq 1. \tag{14}$$

Here, although ϵ_T is not directly observable, we can determine its value by expressing our ARIMA(0,1,1) model in its infinite autoregressive form. The assumption that the invertibility condition $|\theta| < 1$ holds is important at this stage. From Equation (11),

$(1-L)Y_t = (1+\theta L)\epsilon_t$
→ invertibility of θ:
$(1+\theta L)^{-1}(1-L)Y_t = \epsilon_t$
→ recall infinite series for $(1+\theta L)^{-1}$.

$$\begin{aligned} \epsilon_t &= (1 + \theta L)^{-1}(1 - L)Y_t \\ &= (1 - L)\sum_{k=0}^{\infty}(-\theta)^k L^k Y_t \\ &= Y_t - (1 + \theta)\sum_{k=1}^{\infty}(-\theta)^{k-1} L^k Y_t. \end{aligned}$$

the infinite AR form.

$(1+\theta L)^{-1} = \sum_{k=0}^{\infty} (-\theta)^k L^k$

MA(1): $Y_t = \epsilon_t + \theta\epsilon_{t-1}$

$Y_{t-1} = \epsilon_{t-1} + \theta\epsilon_{t-2} \to \epsilon_{t-1} = Y_{t-1} - \theta\epsilon_{t-2}$

$Y_{t-2} = \epsilon_{t-2} + \theta\epsilon_{t-3} \to \epsilon_{t-2} = Y_{t-2} - \theta\epsilon_{t-3}$

$Y_{t-3} = \epsilon_{t-3} + \theta\epsilon_{t-4} \to \epsilon_{t-3} = Y_{t-3} - \theta\epsilon_{t-4}$

$Y_t = \epsilon_t + \theta(Y_{t-1} - \theta\epsilon_{t-2}) = \epsilon_t + \theta Y_{t-1} - \theta^2\epsilon_{t-2} = \epsilon_t + \theta Y_{t-1} - \theta^2(Y_{t-2} - \theta\epsilon_{t-3}) = \epsilon_t + \theta Y_{t-1} - \theta^2 Y_{t-2} + \theta^3\epsilon_{t-3}$

$= \epsilon_t + \theta Y_{t-1} - \theta^2 Y_{t-2} + \theta^3(Y_{t-3} - \theta\epsilon_{t-4}) = \epsilon_t + \theta Y_{t-1} - \theta^2 Y_{t-2} + \theta^3 Y_{t-3} - \theta^4\epsilon_{t-4}$

$Y_t = \varepsilon_t - \sum_{i=1}^{n-1} (-\theta)^i Y_{t-i} - (-\theta)^n \varepsilon_{t-n} = \varepsilon_t - \sum_{i=1}^{n-1} (-\theta)^i L^i Y_t - (-\theta)^n \varepsilon_{t-n}$

Show that $\varepsilon_t - \sum_{i=1}^{n-1} (-\theta)^i Y_{t-i} \xrightarrow{m.s.} Y_t$ as $n \to \infty$ via $E[(\xi - \xi_n)^2] \to 0$ as $n \to \infty$. $\therefore$ show $E[\theta^{2n} \varepsilon_{t-n}^2]$

$= \theta^{2n} \sigma^2 \to 0$ as $n \to \infty$

since $|\theta| < 1$

Thus

$\therefore Y_t = \varepsilon_t - \sum_{i=1}^{\infty} (-\theta)^i L^i Y_t$

$$\epsilon_t = Y_t - (1+\theta) \sum_{k=1}^{\infty} (-\theta)^{k-1} Y_{t-k}, \qquad t \in \mathbb{Z}. \tag{15}$$

Putting $t = T$ in Equation (15) and substituting into Equation (14),

$\varepsilon_t = (1 + \sum_{i=1}^{\infty} (-\theta)^i L^i) Y_t$

$= \sum_{i=0}^{\infty} (-\theta)^i L^i$ *

$$\hat{y}_T(h) = (1+\theta) \sum_{k=0}^{\infty} (-\theta)^k y_{T-k}, \qquad h \geq 1. \tag{16}$$

$Y_t = (1+\theta L)\varepsilon_t \quad \therefore \varepsilon_t = (1+\theta L)^{-1} Y_t$

$\therefore (1+\theta L)^{-1} = (1 + \sum_{i=1}^{\infty} (-\theta)^i L^i)$

In Equation (16), if θ is negative, $\hat{y}_T(h)$ is expressed as an *exponentially weighted moving average* — the coefficients on the right hand side of Equation (16) decrease geometrically, i.e., exponentially, with k.

To be able to use Equation (16) for forecasting, ideally, we should know the history of the process back into the infinite past; but if T is large and/or $|\theta|$ is small, the fact that we do not in practice have observations back into the infinite past has only a negligible effect and we can write

$$\hat{y}_T(h) = (1+\theta) \sum_{k=0}^{T-1} (-\theta)^k y_{T-k}. \tag{17}$$

9.4.2 Alternative derivation

To derive Equation (14) by an alternative method and to obtain in addition an expression for the forecast error, use Equation (12) recursively with $t = T+h, T+h-1, \ldots, T+1$:

$$\begin{aligned} Y_{T+h} &= Y_{T+h-1} + \epsilon_{T+h} + \theta \epsilon_{T+h-1} \\ &= Y_{T+h-2} + \epsilon_{T+h} + (1+\theta)\epsilon_{T+h-1} + \theta \epsilon_{T+h-2} \\ &\cdots \\ &= Y_T + \epsilon_{T+h} + (1+\theta) \sum_{i=1}^{h-1} \epsilon_{T+h-i} + \theta \epsilon_T. \end{aligned}$$

Thus

$$Y_{T+h} = Y_T + \theta \epsilon_T + \epsilon_{T+h} + (1+\theta) \sum_{i=1}^{h-1} \epsilon_{T+h-i}. \tag{18}$$

(A formal proof of the result of Equation (18) may be obtained by induction on h, using Equation (12).)

Taking expectations conditional upon $\mathcal{H}_T$ in Equation (18) and noting that the white noise terms $\epsilon_{T+h}, \epsilon_{T+h-1}, \ldots, \epsilon_{T+1}$ are independent of $\mathcal{H}_T$ with means zero, we re-derive Equation (14).

In addition, from Equations (14) and (18), we find an expression for the forecast error,

$$\begin{aligned} e_T(h) &= Y_{T+h} - \hat{y}_T(h) \\ &= \epsilon_{T+h} + (1+\theta) \sum_{i=1}^{h-1} \epsilon_{T+h-i}. \end{aligned} \tag{19}$$

Thus the first two terms on the right hand side of Equation (18) represent the forecast and the sum of the final two terms is the forecast error. The corresponding forecast error variance is given by

$$\begin{aligned} V(h) &= \text{var}(e_T(h)) \\ &= \text{var}\left(\epsilon_{T+h} + (1+\theta)\sum_{i=1}^{h-1}\epsilon_{T+h-i}\right) \\ &= \left[1+(h-1)(1+\theta)^2\right]\sigma^2. \end{aligned} \tag{20}$$

When $h = 1$, Equations (19) and (20) reduce to

$$e_T(1) = \epsilon_{T+1} \tag{21}$$

and

$$V(1) = \sigma^2,$$

Thus, the white noise terms are the one step ahead forecast errors, whose common variance is the one-step forecast variance.

Note also that $V(h) \uparrow \infty$ as $h \to \infty$, which is a characteristic of integrated models, whereas for stationary models in general $V(h) \uparrow \gamma_0$.

9.5 Forecasting for ARIMA models

Consider the ARIMA(p, d, q) model as specified in Chapter 6,

$$\phi(L)(1-L)^d Y_t = c + \theta(L)\epsilon_t. \tag{22}$$

Assuming that $d \geq 1$, this is an integrated model that does not define a stationary process. Hence the process $\{Y_t\}$ cannot be expressed as an infinite moving average. It is still, however, possible to write, as in Equation (5) of Chapter 8 for the ARMA process,

$$Y_{T+h} = \hat{y}_T(h) + \sum_{i=0}^{h-1}\psi_i\epsilon_{T+h-i}. \tag{23}$$

It turns out that the generating function ψ for the ψ_i is given by

$$\psi(z) = (1-z)^{-d}\phi^{-1}(z)\theta(z), \tag{24}$$

although $\sum_{i=0}^{\infty}\psi_i^2 = \infty$ and so $\sum_{i=0}^{\infty}\psi_i\epsilon_{t-i}$ does not converge, so that an infinite moving average expression for Y_t does not exist. It will always be the case that $\psi_0 = 1$. The $\{\psi_i : i \geq 1\}$ may be determined recursively, or by carrying out the expansion of Equation (24).
For example, for the ARIMA(0,1,1) model, $\psi(z) = (1-z)^{-1}(1+\theta z)$ and $\psi_i = 1+\theta,\ i \geq 1$.
As for the ARMA models, the forecast error is given by

$$e_T(h) = \sum_{i=0}^{h-1}\psi_i\epsilon_{T+h-i} \tag{25}$$

and the forecast error variance by

$$V(h) = \sum_{i=0}^{h-1} \psi_i^2 \sigma^2. \tag{26}$$

In particular, as for the ARMA models, $\psi_0 = 1, e_T(1) = \epsilon_{T+1}$ and $V(1) = \sigma^2$.

Since $\sum_{i=0}^{\infty} \psi_i^2 = \infty$, letting $h \to \infty$ in Equation (26), we find that

$$V(h) \uparrow \infty,$$

in contrast to the ARMA models, for which $V(h)$ converges to a finite limit.

A recursion for the forecast function $\{\hat{y}_T(h) : h \geq 1\}$ is obtained by setting $t = T + h$ in Equation (22) and then taking expectations conditional upon $\mathcal{H}_T$. For $h > \max(p + d, q)$ the recursive relation reduces to

$$\phi(L)(1 - L)^d \hat{y}_T(h) = c, \tag{27}$$

where the lag operator L is applied to h. Equation (27) is a linear difference equation whose auxiliary equation has the root 1 repeated d times, corresponding to the presence of the factor $(1 - L)^d$. Consequently, the general solution of Equation (27), in the homogeneous case $c = 0$, is of the form

$$\hat{y}_T(h) = \sum_{k=1}^{p} A_k \alpha_k^h + \sum_{k=0}^{d-1} B_k h^k. \tag{28}$$

Because $|\alpha_k| < 1,\ 1 \leq k \leq p$, letting $h \to \infty$ in Equation (28), we find that the limiting form of $\hat{y}_T(h)$, the *eventual forecast function*, is given by

$$\hat{y}_T(h) = \sum_{k=0}^{d-1} B_k h^k,$$

a polynomial in h of order $d - 1$.

As mentioned in Chapter 6, it is usual to take $c = 0$ for an ARIMA model with $d \geq 1$, but if we take $c \neq 0$ then Equation (27) is a non-homogeneous equation with particular solution equal to

$$\frac{c}{d!\,(1 - \sum_{k=1}^{p} \phi_k)}\ h^d.$$

In this case the eventual forecast function is given by

$$\hat{y}_T(h) = \sum_{k=0}^{d-1} B_k h^k + \frac{c}{d!\,(1 - \sum_{k=1}^{p} \phi_k)}\ h^d,$$

which is a polynomial in h of order d. The values of the B_k will depend upon $\mathcal{H}_T$, but the final term is deterministic, a function of the parameters which does not depend upon $\mathcal{H}_T$. It is the presence of this inflexible deterministic term that is a principal reason for usually taking $c = 0$ in ARIMA models with $d \geq 1$.

- If necessary, in searching for an appropriate model, rather than taking $c \neq 0$,

  the value of d may be increased by 1, which gives the same degree of polynomial for the eventual forecast function, but with a more flexible form.

9.6 Airline Passengers Example

We found previously that an adequate model for the logarithm of monthly airline passenger numbers between January 1949 and December 1960 is given by ARIMA$(0, 1, 1) \times (0, 1, 1)_{12}$.

With December 1960 as the origin of the forecasts, we will generate the forecasts for the whole of the ensuing decade ("the sixties"): this corresponds to a lead time of $h = 10 \times 12 = 120$ months. We apply the command `predict` to the fitted model stored in `lair.arima`, supplied as its first argument, and 120 as the second argument.

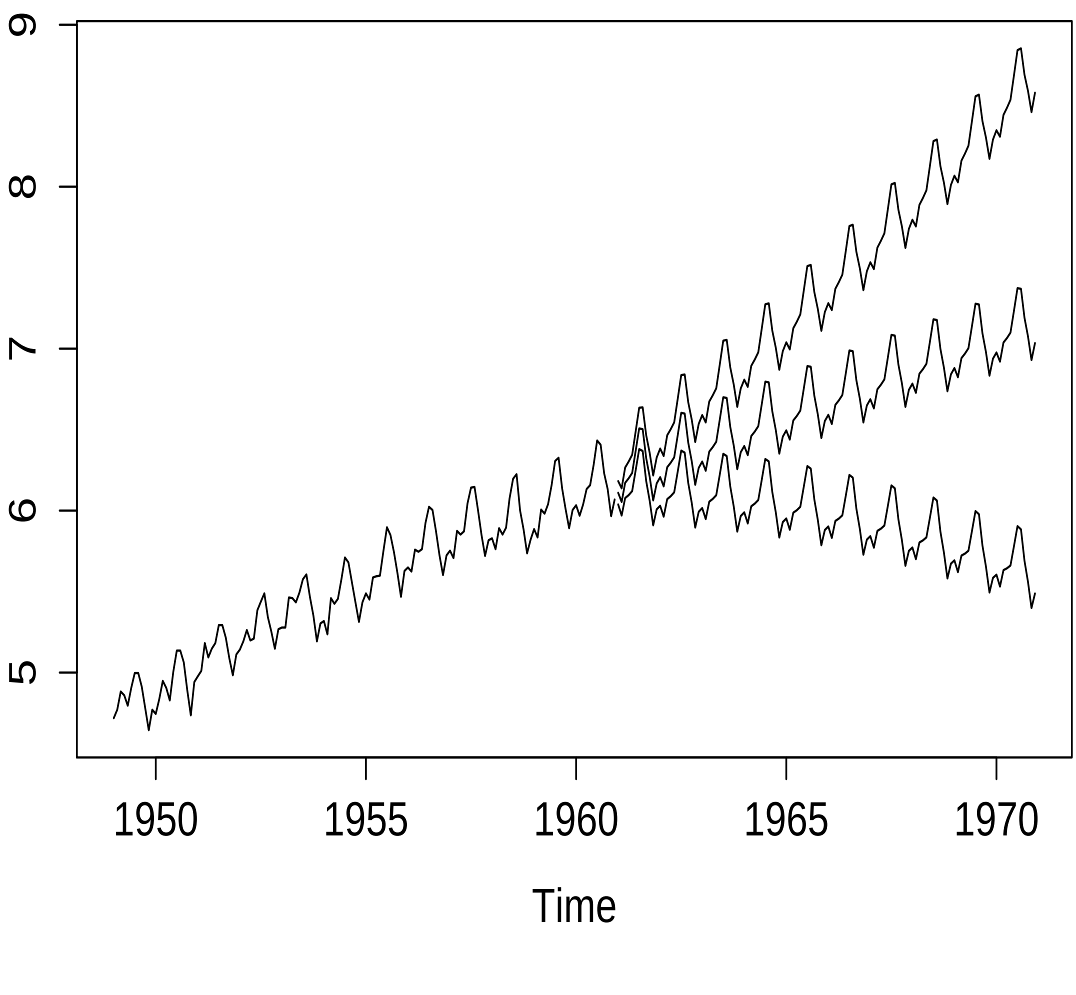
```
> lair.fore <- predict(lair.arima, 120)
> lair.fore
$pred
          Jan      Feb      Mar      Apr      May      Jun      Jul      Aug
1961 6.110186 6.053775 6.171715 6.199300 6.232556 6.368779 6.507294 6.502906
1962 6.206435 6.150025 6.267964 6.295550 6.328805 6.465028 6.603543 6.599156
1963 6.302684 6.246274 6.364213 6.391799 6.425054 6.561277 6.699792 6.695405
1964 6.398933 6.342523 6.460463 6.488048 6.521304 6.657526 6.796042 6.791654
1965 6.495183 6.438772 6.556712 6.584297 6.617553 6.753776 6.892291 6.887903
1966 6.591432 6.535022 6.652961 6.680547 6.713802 6.850025 6.988540 6.984153
1967 6.687681 6.631271 6.749210 6.776796 6.810051 6.946274 7.084789 7.080402
1968 6.783930 6.727520 6.845460 6.873045 6.906301 7.042523 7.181039 7.176651
1969 6.880180 6.823769 6.941709 6.969294 7.002550 7.138773 7.277288 7.272900
1970 6.976429 6.920019 7.037958 7.065544 7.098799 7.235022 7.373537 7.369150
          Sep      Oct      Nov      Dec
1961 6.324698 6.209008 6.063487 6.168025
1962 6.420947 6.305257 6.159737 6.264274
1963 6.517197 6.401507 6.255986 6.360523
1964 6.613446 6.497756 6.352235 6.456773
1965 6.709695 6.594005 6.448484 6.553022
1966 6.805944 6.690254 6.544734 6.649271
1967 6.902194 6.786504 6.640983 6.745520
1968 6.998443 6.882753 6.737232 6.841770
1969 7.094692 6.979002 6.833481 6.938019
1970 7.190941 7.075251 6.929731 7.034268

$se
            Jan        Feb        Mar        Apr        May        Jun
1961 0.03671562 0.04278291 0.04809072 0.05286830 0.05724856 0.06131670
1962 0.09008475 0.09549708 0.10061869 0.10549195 0.11014981 0.11461854
1963 0.14650643 0.15224985 0.15778435 0.16313118 0.16830825 0.17333075
1964 0.20896657 0.21513653 0.22113442 0.22697386 0.23266679 0.23822371
1965 0.27748210 0.28408309 0.29053414 0.29684503 0.30302451 0.30908048
1966 0.35174476 0.35876289 0.36564634 0.37240257 0.37903840 0.38556004
1967 0.43142043 0.43883816 0.44613258 0.45330963 0.46037481 0.46733319
1968 0.51620376 0.52400376 0.53168935 0.53926541 0.54673651 0.55410688
1969 0.60582584 0.61399203 0.62205103 0.63000694 0.63786363 0.64562471
1970 0.70005133 0.70856907 0.71698563 0.72530453 0.73352910 0.74166246
            Jul        Aug        Sep        Oct        Nov        Dec
1961 0.06513124 0.06873441 0.07215787 0.07542612 0.07855851 0.08157070
1962 0.11891946 0.12307018 0.12708540 0.13097758 0.13475740 0.13843405
1963 0.17821177 0.18296261 0.18759318 0.19211216 0.19652727 0.20084534
1964 0.24365393 0.24896574 0.25416656 0.25926308 0.26426132 0.26916676
```

```
1965 0.31502004 0.32084967 0.32657525 0.33220217 0.33773535 0.34317933
1966 0.39197318 0.39828307 0.40449455 0.41061207 0.41663978 0.42258152
1967 0.47418947 0.48094803 0.48761291 0.49418791 0.50067658 0.50708223
1968 0.56138049 0.56856106 0.57565206 0.58265678 0.58957827 0.59641945
1969 0.65329361 0.66087351 0.66836746 0.67577831 0.68310877 0.69036139
1970 0.74970759 0.75766731 0.76554426 0.77334099 0.78105989 0.78870326
> lair.pred <- ts(lair.fore$pred, start = 1961, frequency = 12)
> L95 <- ts(lair.fore$pred - 1.96 * lair.fore$se, start = 1961, frequency = 12)
> U95 <- ts(lair.fore$pred + 1.96 * lair.fore$se, start = 1961, frequency = 12)
> ts.plot(lair.ts, lair.pred, L95, U95)
```

By making use of the information on both $\hat{y}_T(h)$, given by `$pred`, and $\sqrt{V(h)}$, given by `$se`, then we can generate the probability limits (here constructed at 95% - thus we use $z_{0.025} = 1.96$) as per the formula for the $100(1-\alpha)\%$ interval, $\hat{y}_T(h) \pm z_{\alpha/2} \times \sqrt{V(h)}$. Finally we have plotted the original data, the forecasts and their probability limits for easy comparison: this has been facilitated by converting everything to `ts` objects and assigning each of the series to the appropriate start dates.

The probability limits seem to be fairly close to the central forecasts for about a couple of years beyond the origin (i.e. January 1961 to December 1962). However, beyond that the disparity between the probability limits and the forecasts widens to such an extent, that one might question whether forecasts so far into the future can be relied upon. Notice that the trajectories of the probability limits give the impression that they will not converge toward a certain value: this is in line with the calculation of the previous section that $V(h) \to \infty$ as $h \to \infty$.


9
8
7
6
5
1950
1955
1960
1965
1970
Time