**Forecasting – Inference**

It was shown earlier how to write an ARIMA process as an infinite order MA process. To find the standard error of ARIMA forecasts, this needs to be done again.

Example: AR(1) with μ = 0



Then

= E(xn+m|In)

= 

= 

Let m = 1. Then

 = E(xn+1|In)

= 

= 0 + ϕ1wn +  +…

Note that



which is the same value of  shown earlier.

Let m = 2. Then

= E(xn+2|In)

= 

= 

Let m = 3. Then

= E(xn+3|In)

= 

= 

Forecast error: **** = observed - forecast

The variance of the forecast error is: ****

Shumway and Stoffer denote the variance of the forecast error symbolically as .

Approximate (1-α)100% C.I. for xn+m:

****

where Z1-α/2 is the 1 - α/2 quantile of a standard normal distribution.

I refer to this a “confidence interval”. Others may prefer to call this a prediction interval due to adopting the often used naming convention of “prediction intervals” are for random variables and “confidence intervals” are for parameters. In the end, all frequentist-based intervals are confidence intervals. If you want to differentiate between the names, you may. I will just call them confidence intervals or C.I.s for short.

Example: AR(1) with μ = 0

For m = 1:

 = (wn+1 + ϕ1wn + +…)

– (ϕ1wn + +…)

= wn+1

For m = 2:

 = (wn+2 + ϕ1wn+1 + +…)

– ()

= wn+2 + ϕ1wn+1

For m = 3:

 = (wn+3 + ϕ1wn+2 +  +…)

– ()

= wn+3 + ϕ1wn+2 + 

Note that wt ~ independent (0,). Then

**** = Var(wn+1) = 

 = Var(wn+2 + ϕ1wn+1) =  + 

 = Var(wn+3 + ϕ1wn+2 + )

=  +  + = 

In general for an AR(1) process:



The approximate (1-α)100% C.I. for xn+m is

****

Notice what happens to the confidence interval as m increases – it gets WIDER (although not by much for larger values of m). Why does it make sense for the confidence interval to get wider?

In general for an ARIMA(p,d,q), the formulas for the forecast error, variance of the forecast error, and the C.I. can be derived using the infinite order MA representation.

ϕ(B)(1-B)dxt = θ(B)wt can be rewritten as



where ψ(B) = (1+ψ1B+ψ2B2+…) = .

The forecast error is  =  where ψ0 = 1. For example, with an AR(1) and m = 1,  = wn+1 which matches our previous result.

The variance of the forecast error is

****.

The approximate (1-α)100% C.I. used in practice is

****.

Example: MA(1)

The process can be written as xt = (1 + θ1B)wt

Translating this to xt = ψ(B)wt produces ψ1 = θ1 and ψi = 0 for i > 1. Then

****

Example: ARIMA(1,1,1)



Then (1-ϕ1B)(1-B)(1+ψ1B+ψ2B2+ψ3B3+…) = (1+θ1B)

⇔[1-B(1+ϕ1)+ϕ1B2](1+ψ1B+ψ2B2+ψ3B3+…) = (1+θ1B)

⇔1+ψ1B+ψ2B2+ψ3B3+…

-(1+ϕ1)B-ψ1(1+ϕ1)B2-ψ2(1+ϕ1)B3-ψ3(1+ϕ1)B4-…

ϕ1B2+ψ1ϕ1B3+ψ2ϕ1B4+ψ3ϕ1B5 = (1+θ1B)

B: ψ1-(1+ϕ1) = θ1 ⇒ ψ1=1+ϕ1+θ1

B2: ψ2-ψ1(1+ϕ1)+ϕ1 = 0 ⇒ ψ2=ψ1(1+ϕ1)-ϕ1

B3: ψ3-ψ2(1+ϕ1)+ϕ1ψ1 = 0 ⇒ ψ3=ψ2(1+ϕ1)-ϕ1ψ1

Bj:ψj=ψj-1(1+ϕ1)-ϕ1ψj-2

 = 

 = 

Example: AR(1) with ϕ1 = 0.7, μ = 0, and  = 1 (fit\_AR1.R, AR1.0.7.txt)

From the previous R code and output, it was found that  = 1.335638 and  = 0.6853698. The forecasts and standard errors, , are automatically calculated by the predict() function:

> fore.mod <- predict(object = mod.fit, n.ahead = 5, se.fit

 = TRUE)

> fore.mod

$pred

Time Series:

Start = 101

End = 105

Frequency = 1

[1] 1.26014875 0.72767770 0.36273810 0.11261952

[5] -0.05880421

$se

Time Series:

Start = 101

End = 105

Frequency = 1

[1] 1.155698 1.401082 1.502576 1.547956 1.568820

The 95% confidence intervals were found using

**** to produce

> #Calculate 95% C.I.s

> low <- fore.mod$pred - qnorm(p = 0.975, mean = 0, sd =

 1)\*fore.mod$se

> up <- fore.mod$pred + qnorm(p = 0.975, mean = 0, sd =

 1)\*fore.mod$se

> data.frame(low, up)

 low up

1 -1.004978 3.525276

2 -2.018392 3.473748

3 -2.582258 3.307734

4 -2.921319 3.146558

5 -3.133634 3.016026

Doing the calculations by hand produces:

|  |  |  |  |
| --- | --- | --- | --- |
| m |  |  | 95% C.I. |
| 1 | 1.2601 |  = 1.336 | 1.2601±1.96= (-1.005, 3.525) |
| 2 | 0.7277 |  = 1.336 (1+0.68542) = 1.963 | 0.7277±1.96= (-2.018, 3.473) |
| 3 | 0.3627 |  = 1.336 (1+0.68542+0.68544) = 2.258 | 0.3627±1.96= (-2.582, 3.308) |

Forecast plots:

> plot(x = x, ylab = expression(x[t]), xlab = "t", type =

 "o", col = "red", lwd = 1, pch = 20, main =

 expression(paste("Data simulated from AR(1): ", x[t]

 == 0.7\*x[t-1] + w[t], " where ", w[t], "~N(0,1)")) ,

 panel.first = grid(col = "gray", lty = "dotted"), xlim

 = c(1, 105))

> lines(x = c(x - mod.fit$residuals, fore.mod$pred), lwd

 = 1, col = "black", type = "o", pch = 17)

> lines(y = low, x = 101:105, lwd = 1, col = "darkgreen",

 lty = "dashed")

> lines(y = up, x = 101:105, lwd = 1, col = "darkgreen",

 lty = "dashed")

> legend(locator(1), legend = c("Observed", "Forecast",

 "95% C.I."), lty = c("solid", "solid", "dashed"),

 col = c("red", "black", "darkgreen"), pch = c(20,

 17, NA), bty = "n")



> #Zoom in

> plot(x = x, ylab = expression(x[t]), xlab = "t", type =

 "o", col = "red", lwd = 1, pch = 20, main =

 expression(paste("Data simulated from AR(1): ", x[t]

 == 0.7\*x[t-1] + w[t], " where ", w[t], "~N(0,1)")) ,

 panel.first = grid(col = "gray", lty = "dotted"), xlim =

 c(96, 105))

> lines(x = c(x - mod.fit$residuals, fore.mod$pred), lwd

 = 1, col = "black", type = "o", pch = 17)

> lines(y = low, x = 101:105, lwd = 1, col = "darkgreen",

 lty = "dashed")

> lines(y = up, x = 101:105, lwd = 1, col = "darkgreen",

 lty = "dashed")

> legend(locator(1), legend = c("Observed", "Forecast",

 "95% C.I."), lty = c("solid", "solid", "dashed"), col

 = c("red", "black", "darkgreen"), pch = c(20, 17,

 NA), bty = "n")



Example: ARIMA(1,1,1) with ϕ1 = 0.7, θ1 = 0.4,  = 9, n = 200 (arima111\_sim.R, arima111.csv)

From the previous R code and output:

> mod.fit <- arima(x = x, order = c(1, 1, 1))

> mod.fit

Call:

arima(x = x, order = c(1, 1, 1))

Coefficients:

 ar1 ma1

 0.6720 0.4681

s.e. 0.0637 0.0904

sigma^2 estimated as 9.558: log likelihood = -507.68, aic = 1021.36

> fore.mod <- predict(object = mod.fit, n.ahead = 5, se.fit

 = TRUE)

> fore.mod

$pred

Time Series:

Start = 201

End = 205

Frequency = 1

[1] -486.3614 -484.9361 -483.9784 -483.3348 -482.9023

$se

Time Series:

Start = 201

End = 205

Frequency = 1

[1] 3.091673 7.303206 11.578890 15.682551 19.534208

> #Calculate 95% C.I.s

> low <- fore.mod$pred - qnorm(p = 0.975, mean = 0, sd =

 1)\*fore.mod$se

> up <- fore.mod$pred + qnorm(p = 0.975, mean = 0, sd =

 1)\*fore.mod$se

> data.frame(low, up)

 low up

1 -492.4209 -480.3018

2 -499.2502 -470.6221

3 -506.6726 -461.2842

4 -514.0720 -452.5976

5 -521.1886 -444.6159

Doing the calculations by hand produces:

|  |  |  |  |
| --- | --- | --- | --- |
| m |  |  | 95% C.I. |
| 1 | -486.36 |  | -486.36±1.96= (-492.42, -480.30) |
| 2 | -484.94 | =53.334 | -484.94±1.96= (-499.25, -470.62) |

Because the data is being simulated from an ARIMA(1,1,1) process, 210 observations were actually simulated originally so that one could determine how good the forecasts are for the last 10 time points. Below are the new observations with their forecasts.

| m |  | 95% C.I. lower limit | 95% C.I. upper limit | X200+m |
| --- | --- | --- | --- | --- |
| 1 | -486.36 | -492.42 | -480.30 | -494.85 |
| 2 | -484.94 | -499.25 | -470.62 | -506.44 |
| 3 | -483.98 | -506.67 | -461.28 | -517.70 |
| 4 | -483.33 | -514.07 | -452.60 | -526.64 |
| 5 | -482.90 | -521.19 | -444.62 | -529.10 |
| 6 | -482.61 | -527.92 | -437.30 | -530.94 |
| 7 | -482.42 | -534.26 | -430.58 | -532.52 |
| 8 | -482.29 | -540.20 | -424.37 | -537.46 |
| 9 | -482.20 | -545.78 | -418.61 | -540.39 |
| 10 | -482.14 | -551.04 | -413.24 | -541.43 |

Below is plot of these values. See the program for the code.



Notice that most of the 10 extra observations are outside of the C.I. bounds. Also, notice that the width of the C.I. increases as the time from t = 200 increases.

A few other realizations were simulated to see if the above was an anomaly. The forecast C.I.s did contain the 10 extra observations!

All forecasts and associated inferences assume that what has occurred in the past will occur in the future. Thus, trends seen in the past will continue in the future. What could happen if this does not occur?