Forecasting

Forecasting Is Extrapolating!

First, note that:

Forecasting is ***extrapolating***

STAT 5371 Example

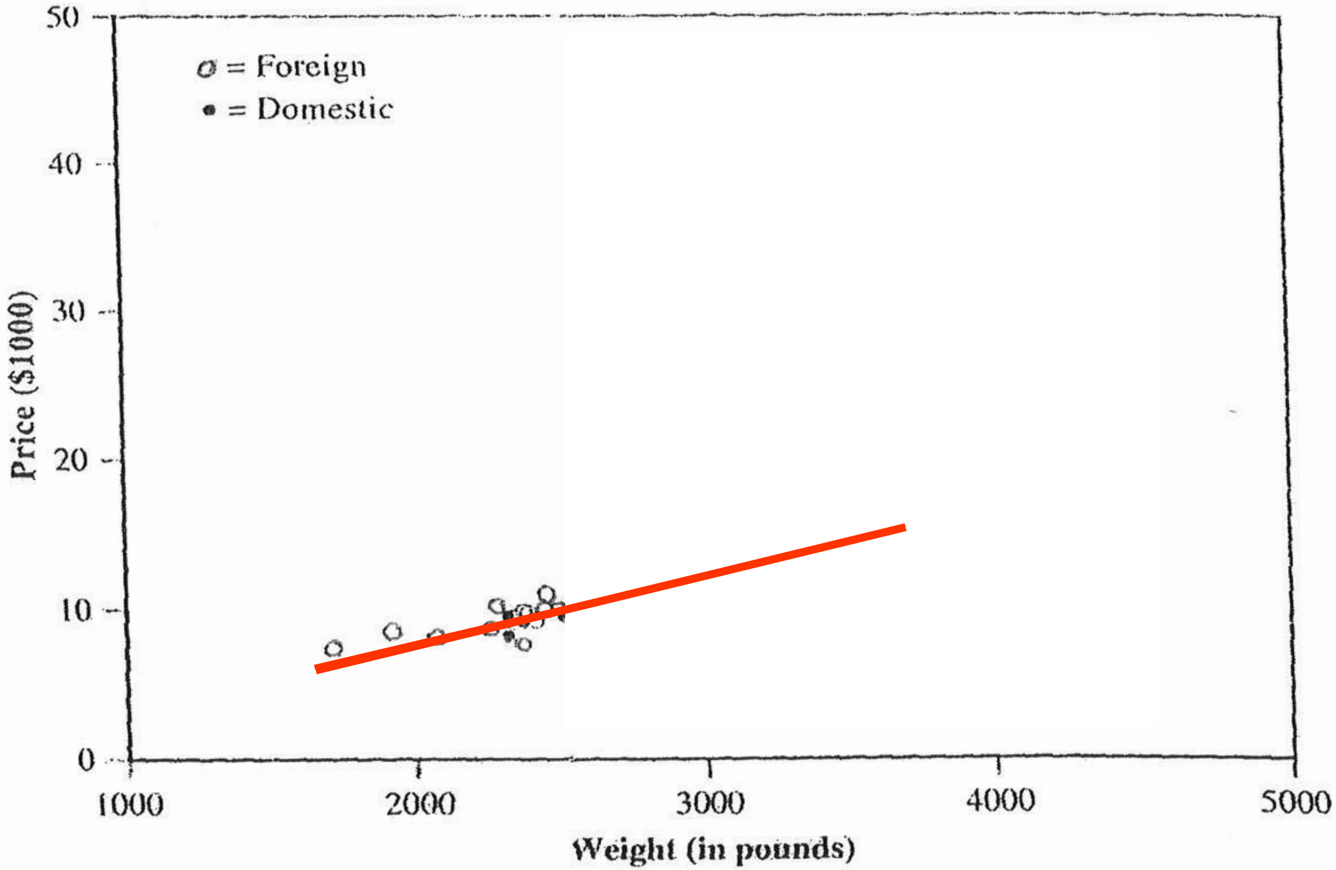


FIGURE 2.9 Base price in thousands of dollars versus weight in pounds for 1991 four-door sedans (Exercise 2.7).

Predict the price of a car that weighs 3500 pounds.

- Extrapolation would say that it's about $16,000.

STAT 5371 Example

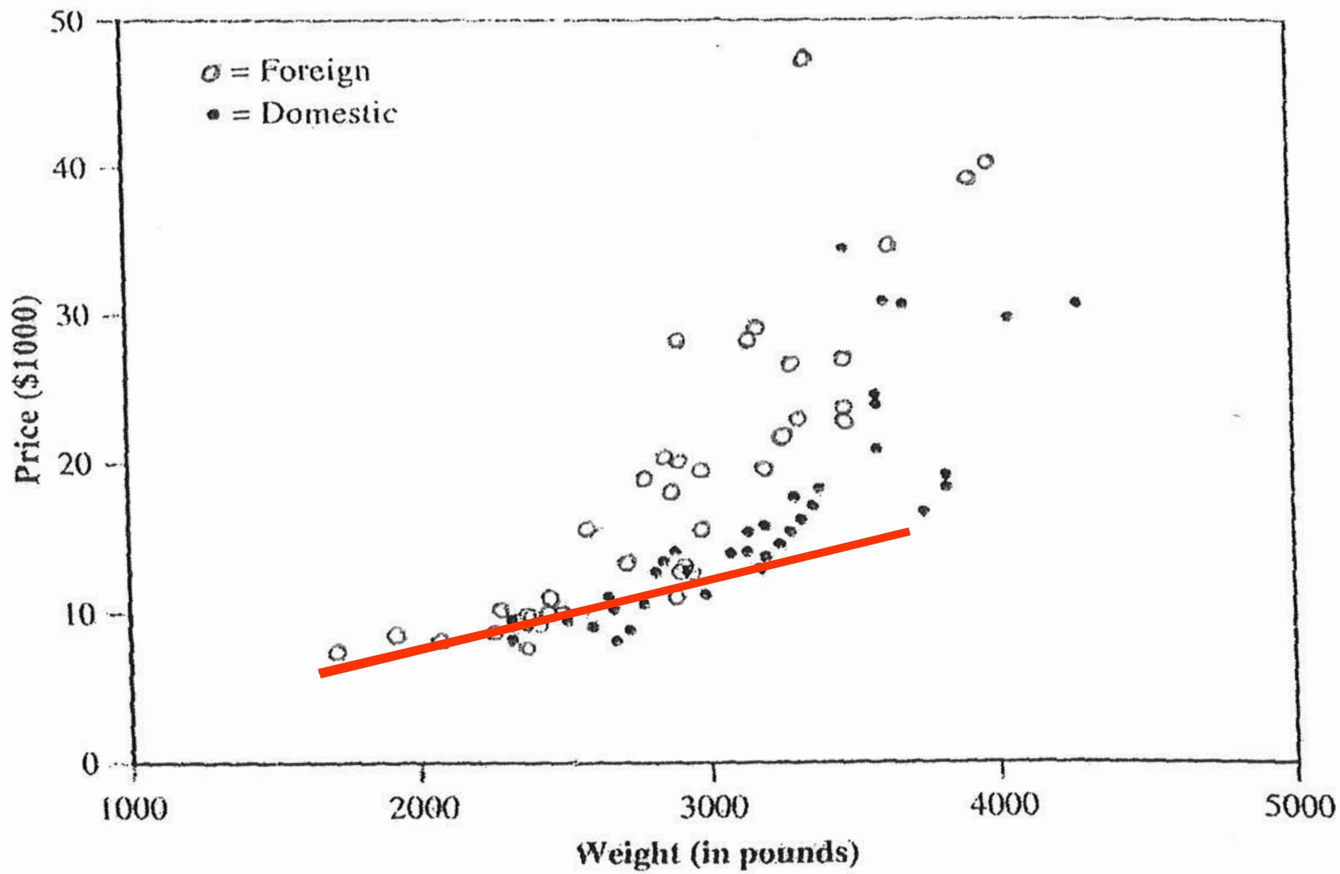


FIGURE 2.9 Base price in thousands of dollars versus weight in pounds for 1991 four-door sedans (Exercise 2.7).

Predict the price of a car that weighs 3500 pounds.

- Extrapolation would say that it's about $16,000.

Oops!

Note: Forecasting is extrapolating.

Lesson from regression analysis and 5371 example?

- ***Be careful when extrapolating!***

However, it’s not very important to:

- “Predict” the sunspot number for 2012
- “Predict” sales for the previous quarter

...

Sometimes we need to extrapolate!

That is, sometimes we need to forecast!

DataScience@SMU

Various Forecasts

Deterministic Signal-Plus-Noise Models

$$\boxed{X_t = s_t + Z_t}$$

s_t is a deterministic signal

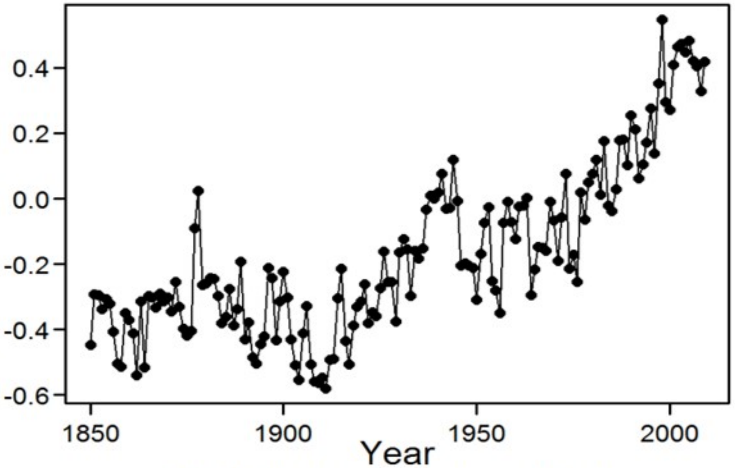
Z_t is a zero-mean, stationary process

Example signals: $s_t = a + bt$

$$s_t = a + bt + ct^2$$

$s_t = A\cos(2\pi ft + C)$, C constant

Recall: Sometimes it’s not easy to tell whether a deterministic signal is present in the data

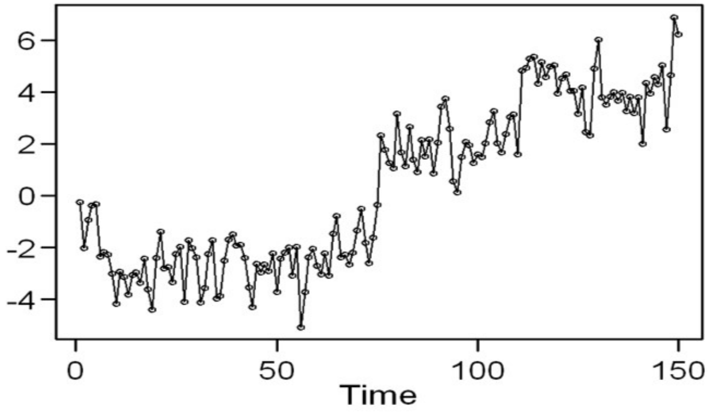


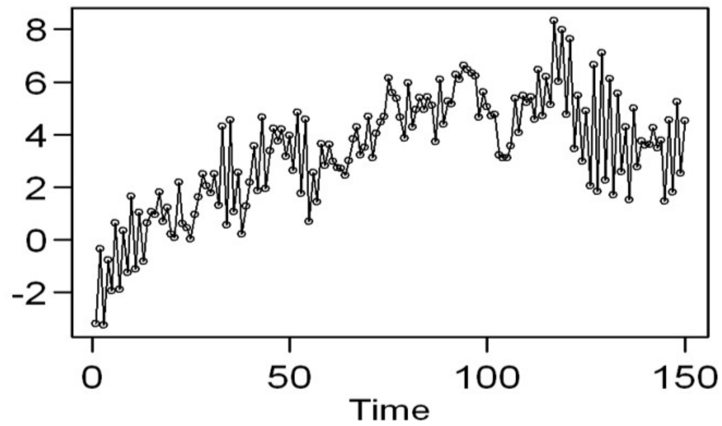
Global Temperature Data

Is there a deterministic signal?

Realizations:

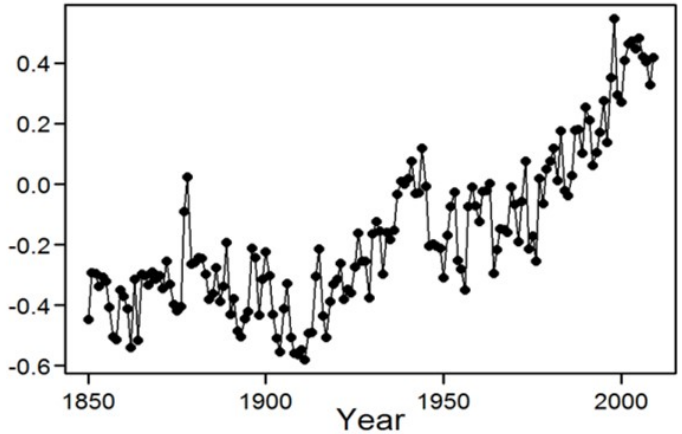
- Is there a deterministic signal?





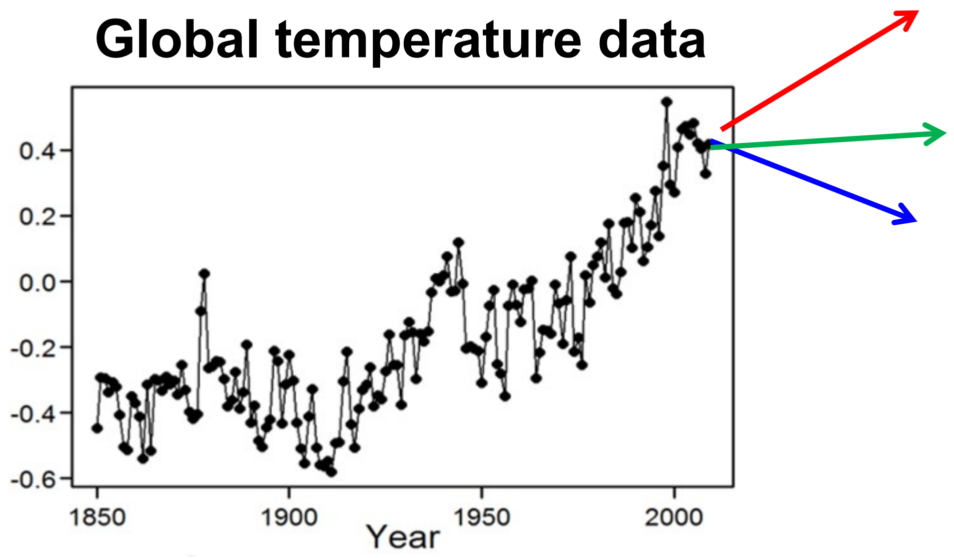
Recall: Sometimes it's not easy to tell whether a deterministic signal is present in the data.

Global temperature data



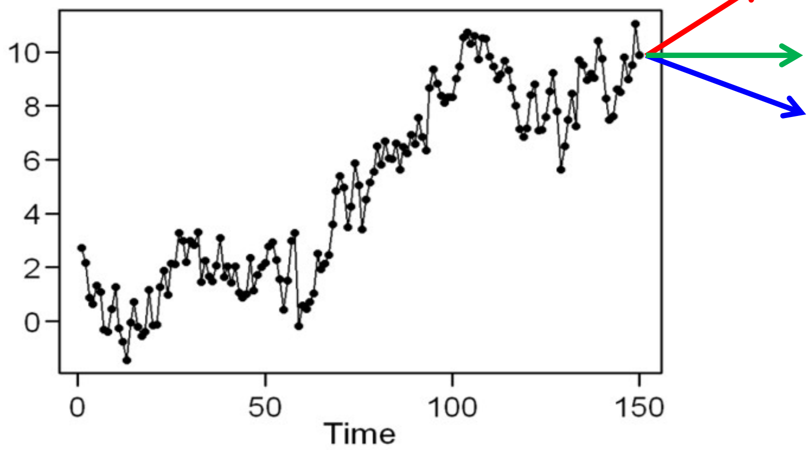
How Would You Predict/Forecast into the Future?

It depends on the model.

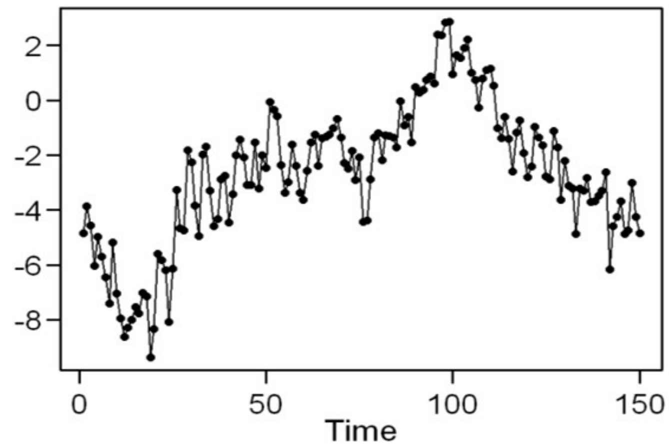


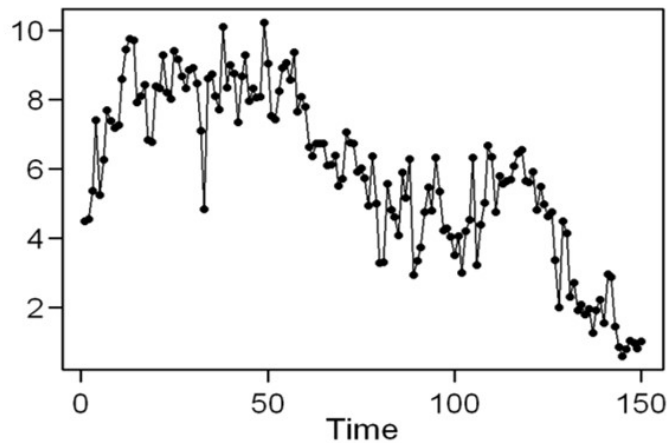
Suppose I told you that this is a realization from the stationary model

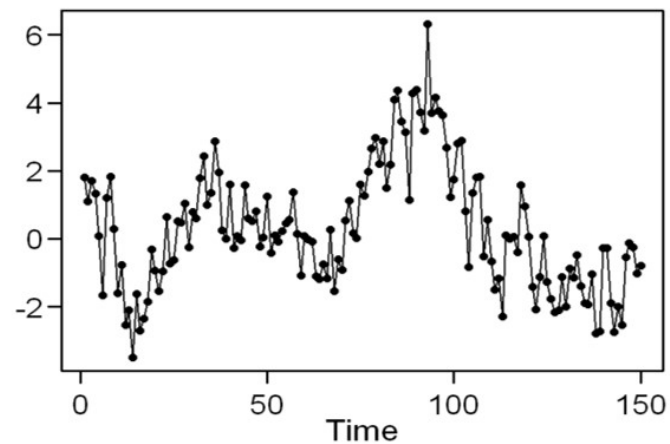
$$(1 \quad .99B)(1 \quad .2B + .455B^2)(1 \quad .53B)X_t = a_t$$



Other realizations (from this model)







Now which forecast do you choose?

DataScience@SMU

Forecasting Setting

Forecasting Setting in this Chapter

Forecast future behavior of a time series given a finite realization of its past.

- The use of ARMA models for this purpose has become popular in recent years.

ARMA forecasting vs. curve fitting

- Curve fitting (regression)
  - Underlying assumption that future behavior follows some deterministic path with only random fluctuations
- ARMA and other time series-based forecasting
  - Underlying assumption is that the future is guided only by its correlation to the past

Box-Jenkins Approach

G. E. P. Box and G. M. Jenkins popularized the use of ARMA models for forecasting.

- The most recent version of their classic text is Box, Jenkins, and Reinsel (2008).
- We will utilize their approach for obtaining ARMA, ARIMA, and seasonal forecasts.
  - In lecture and in **tswge**

Important

In this unit, we will assume that we know the true model (i.e., p, q, the φ's and θ's).

DataScience@SMU

Strategy and Notation

Strategy and Notation

Time series data are available: $X_1, X_2, ..., X_{t_0}$

Goal is to forecast a future (usually unknown) value, say X_{t_0+4}

Note:

The typical case is $t_0 = n$, i.e. $X_{t_0} = X_n$ is the last value in the observed realization

- however, we may want to forecast, say, the last 4 values in the realization (which are known values) to assess the performance of our forecasts
  - in this case $t_0 = n \quad 4$

Strategy and Notation

$\hat{X}_{t_0}(\ell)$ is the forecast of $X_{t_0+\ell}$ given data up to time t_0

- t_0 is called the ***forecast origin***
- ℓ is the ***lead time***, i.e. the number of time units (***steps ahead***) which we want to forecast

For example, suppose we observe $X_1, X_2, \ldots, X_{10}$ $(= X_{t_0})$. and we want to forecast X_{12} $(= X_{t_0+2})$.

Then $\hat{X}_{10}(2)$ is the forecast of X_{12}

And $\hat{X}_{10}(3)$ is the forecast of X_{13}

And $\hat{X}_{10}(7)$ is the forecast of X_{17}

And $\hat{X}_{10}(l)$ is the forecast of X_{10+l}

Strategy and Notation

$\boxed{\hat{X}_{t_0}(\ell) \text{ is the forecast of } X_{t_0+\ell} \text{ given data up to time } t_0}$

- $\hat{X}_{t_0}(\ell) = X_{t_0+\ell}$ if $\ell \le 0$

For example, suppose we observe $X_1, X_2, \ldots, X_{10}\ (= X_{t_0})$.

$\hat{X}_{10}(\ell)$ is a forecast of $X_{10+\ell}$, so

$\hat{X}_{10}(0)$ is a forecast of $X_{10+0} = X_{10}$, which has ***already been observed.***

It makes intuitive sense that $\hat{X}_{10}(0) = X_{10}$

$$\hat{X}_{10}(\quad 1) = X_9$$

$$\hat{X}_{10}(\quad 2) = X_8$$

And so on

DataScience@SMU

Forecasting with AR(1)

Forecasting Using an AR(1) Model

Suppose that X_t is AR(1) (i.e., $X_t - \varphi_1 X_{t-1} = (1-\varphi_1)\mu + a_t$).

And suppose again we observe $X_1, X_2, \ldots, X_{10} \;\; (= X_{t_0})$.

So, $X_{10+2} = \;\; {}_1X_{10+1} + a_{10+2} + \;\; (1 \quad\;\; {}_1)$ **Light board**

i.e., $X_{12} = \;\; {}_1X_{11} + a_{12} + \;\; (1 \quad\;\; {}_1)$

Notes:

- recall that we are assuming we know φ_1
- however, we still can’t use this formula to calculate X_{12} because we don’t know:
    - X_{11}, a_{12}, and μ
- we will estimate μ with $\bar{X}$

General Forecasting Formula for an AR(1)

$$\boxed{\hat{X}_{t_0}(\ell) = \varphi_1 \hat{X}_{t_0}(\ell - 1) + \bar{X}(1 - \varphi_1)}$$

Light board

Note: This formula is used recursively.
That is, we iteratively calculate:

$$\hat{X}_{t_0}(1)$$

$$\hat{X}_{t_0}(2)$$

$$\vdots$$

$$\hat{X}_{t_0}(\ell - 1)$$

$$\boxed{\hat{X}_{t_0}(\ell)}$$

Forecasting Using an AR(1) Model (Example)

We observe $X_1, X_2, \ldots, X_{10} \, (= X_{t_0})$ and want to forecast X_{12}

$$X_{12} = \varphi_1 X_{11} + a_{12} + \bar{X}(1-\varphi_1)$$

Light board

Based on data to time $t = 10$, we ***still*** can't calculate X_{12} because we don't know X_{11} or a_{12}.

Strategy: We will need to ***forecast X_{12} iteratively*** (based on data to time $t = 10$).

First: We forecast X_{11} based on data to time $t = 10$.

X_{11} is given by $X_{11} = \varphi_1 X_{10} + a_{11} + \bar{X}(1-\varphi_1)$

replacing actual (unknown) values with forecasts as of $t = 10$

$$\hat{X}_{10}(1) = \varphi_1 \hat{X}_{10}(0) + \hat{a}_{10}(1) + \bar{X}(1-\varphi_1)$$

Forecasting Using an AR(1) Model (Example)

Light board

$$\hat{X}_{10}(1) = \varphi_1 \hat{X}_{10}(0) + \hat{a}_{10}(1) + \bar{X}(1-\varphi_1)$$

Recall: $\hat{X}_{10}(0) = X_{10}$ (known value)

How about: $\hat{a}_{10}(1)$?

a_t is random white noise with zero mean, and as of $t = 10$ we do not know a_{11} -- but we do know it has zero mean.

So, it makes intuitive sense to define $\hat{a}_{10}(1) = 0,$
and correspondingly, for $\ell > 0,$ we define $\hat{a}_{10}(\ell) = 0$

Now: $\hat{X}_{10}(1) = \varphi_1 X_{10} + \bar{X}(1-\varphi_1)$ (which we can calculate)

And finally: $\hat{X}_{10}(2) = \varphi_1 \hat{X}_{10}(1) + \hat{a}_{10}(2) + \bar{X}(1-\varphi_1)$

$= \varphi_1 \hat{X}_{10}(1) + \bar{X}(1-\varphi_1)$ (which again we can calculate)

Example: AR(1) forecasts

Suppose that the realization $X_1, \ldots, X_{80}$ is observed from the AR(1) model $(1 - .8B)(X_t - 25) = a_t$ and further that $X_{80} = 21.77$ and $\bar{X} = 24.17$. (These data are in tswge file fig6.1nf.)

Although we know $\mu = 25$, we use $\bar{X} = 24.17$ since in practice, μ will be unknown to us.

Calculate $\hat{X}_{80}(\ell), \ell = 1, 2, 3$

Light board

Forecast Function :

$$\hat{X}_{80}(\ell) = .8\hat{X}_{80}(\ell - 1) + \bar{X}(1 - .8) = .8\hat{X}_{80}(\ell - 1) + \underbrace{24.17(.2)}_{4.83}$$

So, $\hat{X}_{80}(1) = .8\hat{X}_{80}(0) + 4.83 = 17.42 + 4.83 = 22.25$

$$\hat{X}_{80}(2) = .8\hat{X}_{80}(1) + 4.83 = .8(22.25) + 4.83 = 22.63$$

$$\hat{X}_{80}(3) = .8\hat{X}_{80}(2) + 4.83 = .8(22.63) + 4.83 = 22.93$$

DataScience@SMU

Example

Forecasting with an AR(1) Model

Example: AR(1) forecasts

$\hat{X}_{80}(1) = 22.25$

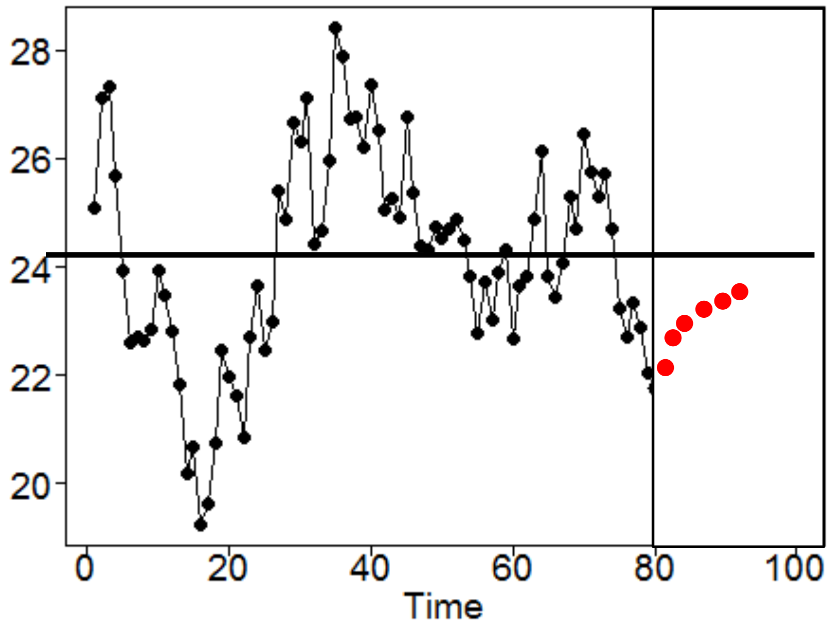
$\hat{X}_{80}(2) = 22.63$

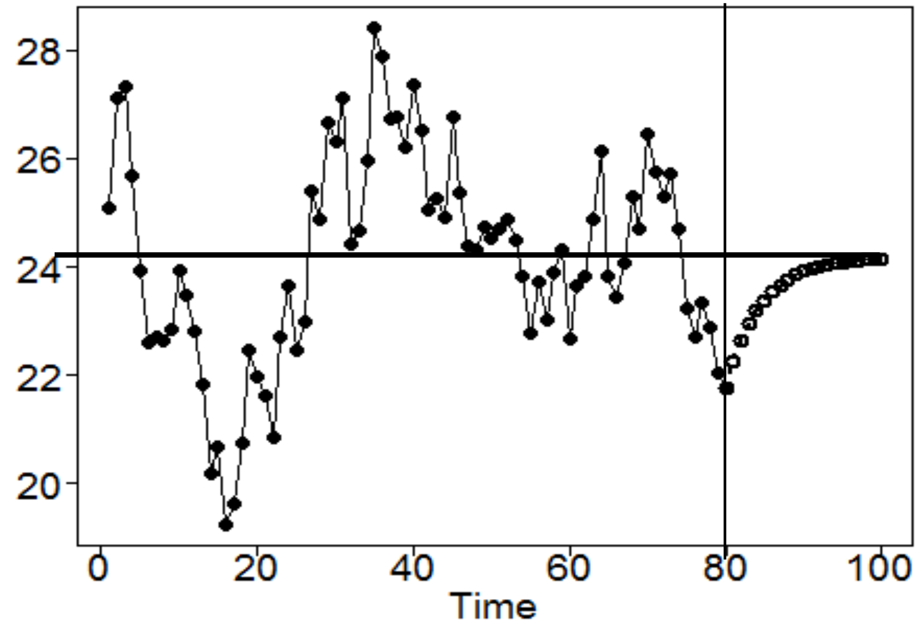
$\hat{X}_{80}(3) = 22.93$

What is $\hat{X}_{80}(4)$? 23.17

What is $\hat{X}_{80}(5)$? 23.37

What is $\hat{X}_{80}(6)$? 23.52





DataScience@SMU

Forecasting with AR(2) and AR(p)

Basic Formula for AR(p) Forecasts

$$\boxed{\begin{aligned}\hat{X}_{t_0}(\ell) = \;\; {}_1\hat{X}_{t_0}(\ell \quad 1) + \cdots + \;\; {}_p\hat{X}_{t_0}(\ell \quad p) \\ + \bar{X}(1 \quad {}_1 \quad \cdots \quad {}_p)\end{aligned}}$$

Example: AR(2)

$$\boxed{\hat{X}_{t_0}(\ell) = \varphi_1 \hat{X}_{t_0}(\ell - 1) + \varphi_2 \hat{X}_{t_0}(\ell - 2) + \bar{X}(1 - \varphi_1 - \varphi_2)}$$

$X_1, X_2, \ldots, X_{74}, X_{75}$ is a realization from the AR(2) model

$(1 \quad 1.6B + .8B^2)(X_t \quad 30)$ where $\bar{X} = 29.4,\ X_{74} = 27.7,\ X_{75} = 23.4$

Forecasts

Light board

$$\hat{X}_{75}(\ell) = 1.6\hat{X}_{75}(\ell - 1) - .8\hat{X}_{75}(\ell - 2) + 29.4(1 - 1.6 + .8)$$

$$\hat{X}_{75}(1) = 1.6\hat{X}_{75}(0) - .8\hat{X}_{75}(-1) + 5.9$$

$$\hat{X}_{75}(1) = 1.6X_{75} - .8X_{74} + 5.9 = 1.6(23.4) - .8(27.7) + 5.9 = 21.2$$

$$\hat{X}_{75}(2) = 1.6\hat{X}_{75}(1) - .8X_{75} + 5.8 = 1.6(21.2) - .8(23.4) + 5.9 = 21.1$$

$$\hat{X}_{75}(3) = 1.6(21.1) \quad .8(21.2) + 5.9 = 22.7$$

DataScience@SMU

Example

Forecasting with an AR(2) Model

Example: AR(2) forecasts

$\hat{X}_{75}(1) = 21.2$

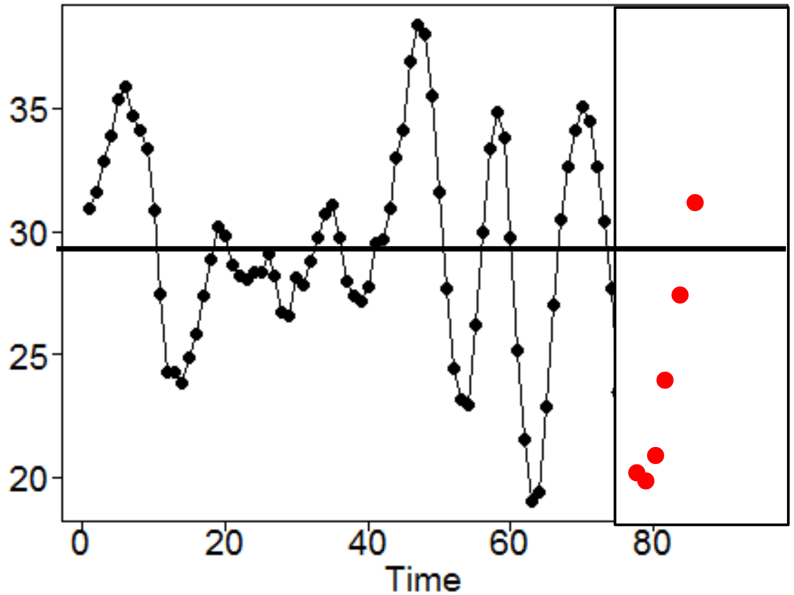
What is $\hat{X}_{75}(4)$? 25.3

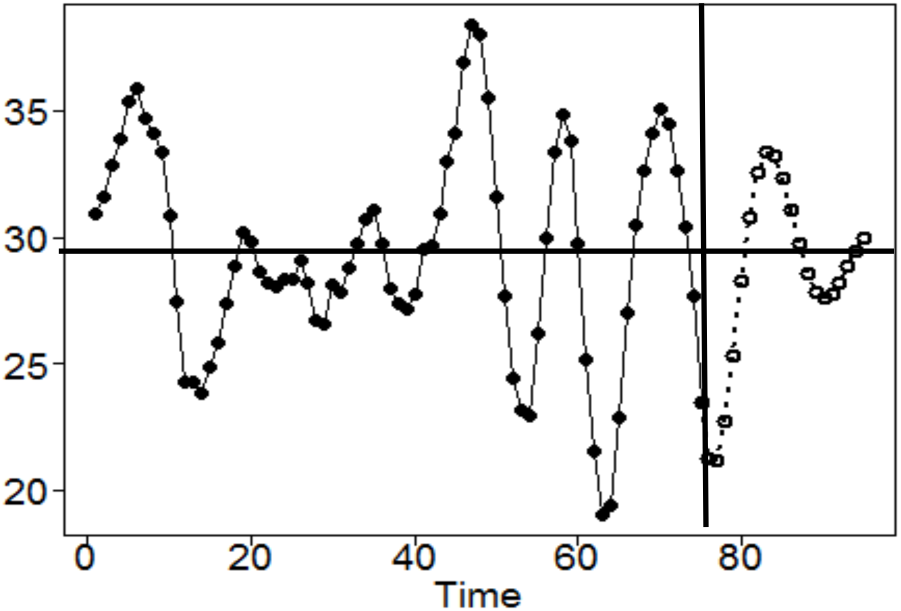
What is $\hat{X}_{75}(5)$? 28.2

$\hat{X}_{75}(2) = 21.1$

What is $\hat{X}_{75}(6)$? 30.8

$\hat{X}_{75}(3) = 22.7$





DataScience@SMU

tswge

Forecasting AR(*p*) Models

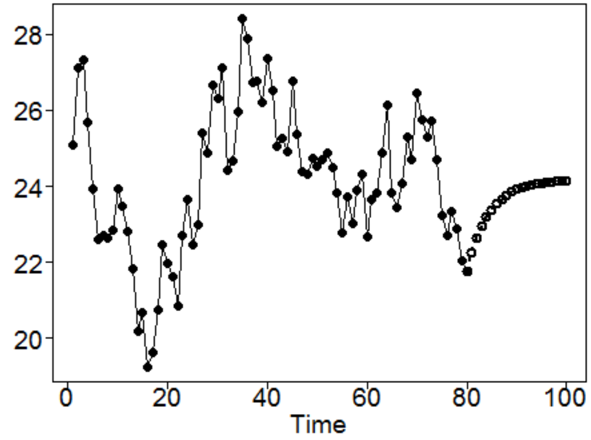
tswge Forecasting Examples

AR(1) example (revisited)

Realization $X_1, \ldots , X_{80}$ in tswge file fig6.1nf from the AR(1) model
$(1 - .8B)(X_t - 25) = a_t$

```
data(fig6.1nf)
fore.arma.wge(fig6.1nf,phi=.8,n.ahead=20,plot=TRUE,limits=FALSE)
```

These commands plot the realization and the n.ahead=20 forecasts shown below. The values of the forecasts are in $f of the output.



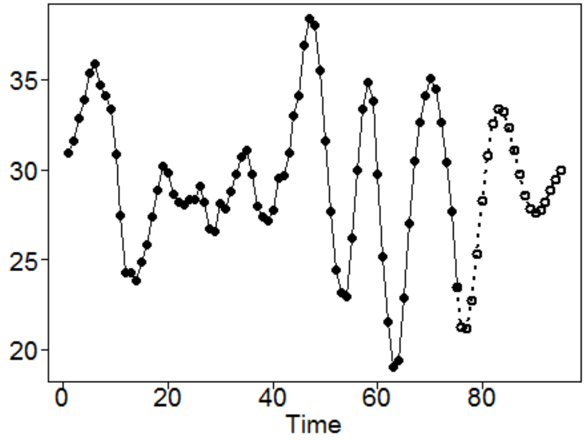
tswge Forecasting Examples

AR(2) example (revisited)

Realization $X_1,\ldots,X_{75}$ from the AR(2) model $(1 - 1.6B + .8B^2)(X_t - 25) = a_t$

```
x2=gen.arma.wge(n=75,phi=c(1.6,-.8),sn=24)
x2=x2+25
fore.arma.wge(x2,phi=c(1.6,-.8),n.ahead=20,limits=FALSE)
```

These commands plot the realization and the n.ahead=20 forecasts shown below. The values of the forecasts are in $f of the output.



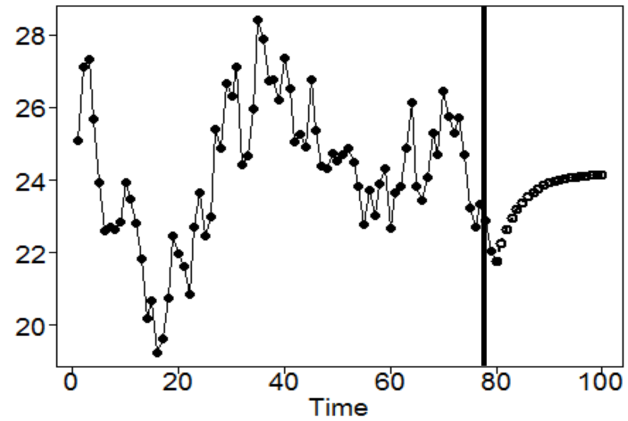
DataScience@SMU

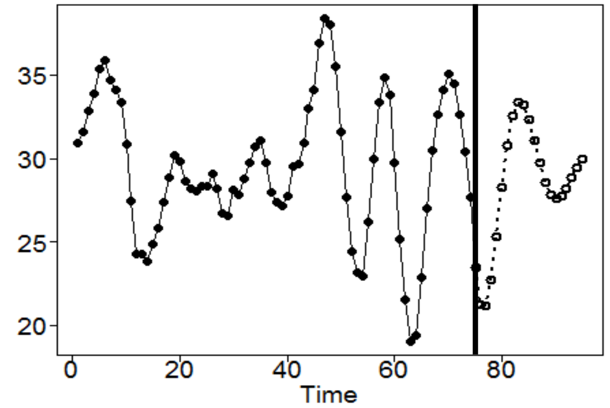
Eventual Forecast Function

Eventual Forecast Function

Two key observations from two previous forecast plots obtained using:

1. In both cases, the forecasts behave like the autocorrelations.
   - AR(1) case: forecasts have exponential shape (although increasing instead of decreasing)
   - AR(2) case: forecasts look like a damped sinusoid (AR part has complex conjugate roots)





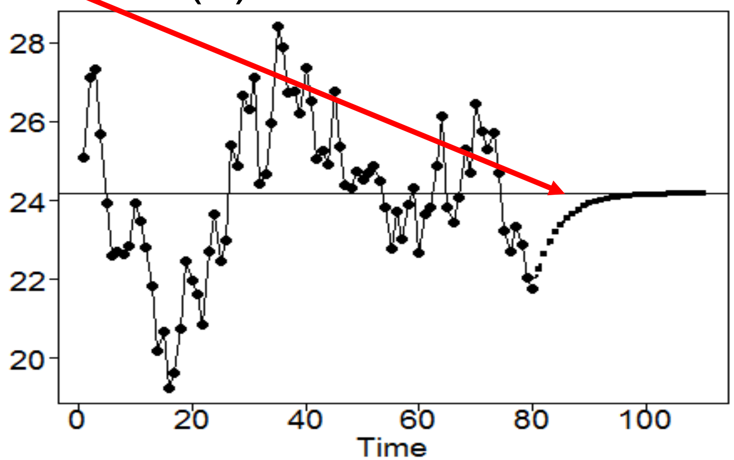
Eventual Forecast Function

2. The forecasts ***tend toward the realization mean*** as number of steps ahead increases.
   - The correlation between observed values (up to time t_0) and $X_{t_0+\ell}$ gets small as ℓ gets large
   - Since the observed values give very little information about $X_{t_0+\ell}$, so it is reasonable to simply forecast $\bar{X}$

Forecasts as ℓ gets large

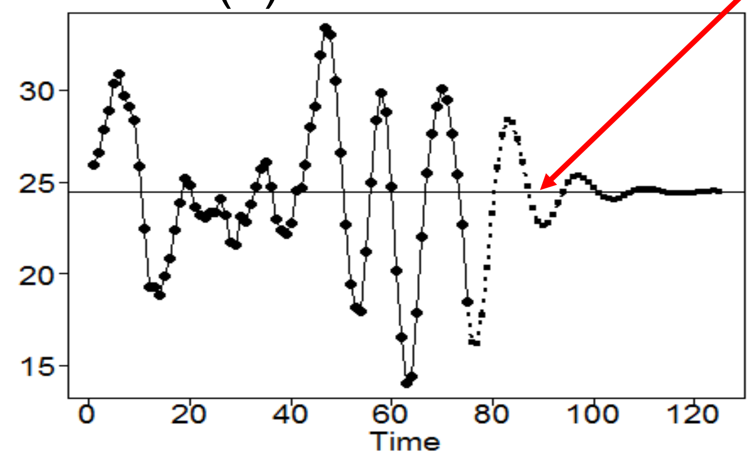
$\bar{X} = 24.17$

AR(1) with 30 forecasts



$\bar{X} = 24.44$

AR(2) with 50 forecasts



DataScience@SMU

Basic Formula for Forecasting with ARMA(p,q)

Basic Formula for ARMA(p,q) Forecasts

$$\hat{X}_{t_0}(\ell) = \sum_{j=1}^{p} \varphi_j \hat{X}_{t_0}(\ell - j) - \sum_{j=\ell}^{q} \theta_j \hat{a}_{t_0+\ell-j} + \bar{X}\left[1 - \sum_{i=1}^{p} \varphi_i\right]$$

Note: These forecasts depend on white-noise variance estimates. Obtaining these estimates is computationally intensive. (See Woodward et al. (2017).)

- The MA part of the model affects the forecasts, and the eventual forecasts behave in much the same way as they would for the AR part of the model.

In practice, we will obtain ARMA(p,q) (and AR(p)) forecasts using **tswge** function **fore.arma.wge.**

DataScience@SMU

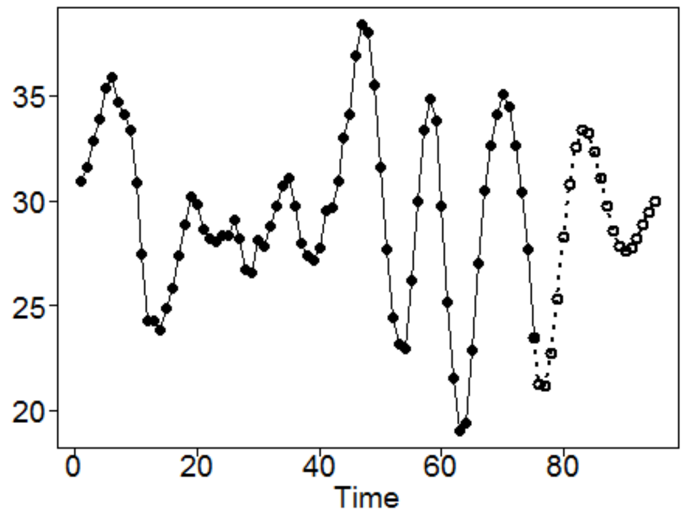
tswge

ARMA(p,q) Forecast Examples

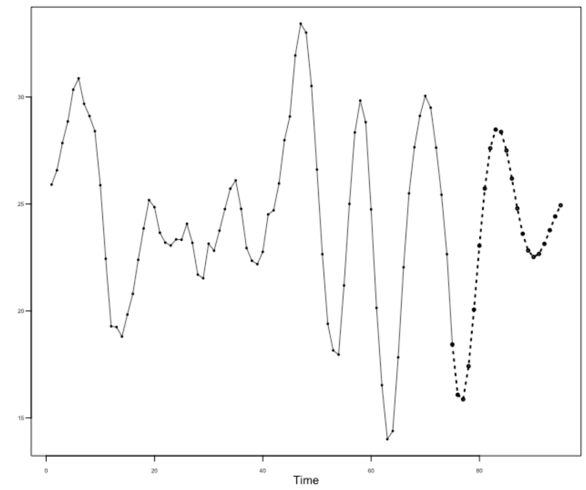
tswge Forecasting Examples

ARMA(2,1) example in Woodward et al. (2017)

```
x2=gen.arma.wge(n=75,phi=c(1.6,-.8),sn=24)
x2=x2+25
fore.arma.wge(x2,phi=c(1.6,-.8), n.ahead=20,limits=FALSE)
```




```
fore.arma.wge(x2,phi=c(1.6,-.8), theta = -.9, n.ahead=20,limits=FALSE)
```



See Woodward et al. (2017) Appendix 6B for more details.

DataScience@SMU

tswge

Canadian Lynx Example

tswge Forecasting Examples: Canadian Lynx

Show how increasing the order of the AR factor helps grab the cycle in the data AR(4), ARMA(4,1), and AR(11).

$$(1 - 1.3B + 0.7B^2 - 0.1B^3 + 0.2B^4)(X_t - 2.9) = a_t$$

AR(4): phi = c(1.3, –0.7, 0.1, –0.2)

$$(1 - 0.7B - 0.1B^2 + 0.2B^3 + 0.3B^4)\,(X_t - 2.9) = (1 + .6B)a_t$$

ARMA(4,1): phi = c(0.7, 0.1, –0.2, –0.3), theta = –.6

AR(11)
phi = c(1.1676, –0.5446, 0.2662, –0.3094, 0.1540, –0.1463, 0.0569, –0.0294, 0.1346, 0.2021, –0.3394)

DataScience@SMU

Probability Limits and ARMA Forecasts

Setting

Probability Limits for ARMA Forecasts

Setting

Given $\hat{X}_{t_0}(\ell)$, an estimate of $X_{t_0+\ell}$ based on data to time t_0, we want to construct limits such that the probability that $X_{t_0+\ell}$ will fall within the limits is 95%.

Recall

An ARMA(p,q) model can be expressed as a General Linear Process (GLP) given by

$$X_t \quad = \quad {}_{j=0} \quad {}_j a_{t \quad j} \text{ where } \quad {}_0 = 1$$

and a_t is normal white noise (mean 0 and variance $\ {}^2_a$)

Probability Limits for ARMA Forecasts

Facts

A. $e_{t_0}(\ell) = X_{t_0+\ell} - \hat{X}_{t_0}(\ell)$ (forecast error)

$$= \sum_{k=0}^{\ell-1} \psi_k a_{t_0+\ell-k}$$

B. $e_{t_0}(\ell)$ is normally distributed (linear combination of normals)

C. $E\left[e_{t_0}(\ell)\right] = 0$ (i.e. $E[X_{t_0+\ell} - \hat{X}_{t_0}(\ell)] = 0$)

D. $\operatorname{var}\left(e_{t_0}(\ell)\right) = \operatorname{var}\left(\sum_{j=0}^{\ell-1} \psi_j a_{t_0+\ell-j}\right)$

$$= \sigma_a^2 \sum_{j=0}^{\ell-1} \psi_j^2$$

DataScience@SMU

Psi Weights and ARMA Models

A Note About the ψ_j's for ARMA(p,q) Models

$(B)X_t = \quad (B)a_t$ can be expressed as

$$X_t = \quad ^{1}(B) \ \ (B)a_t$$

$$= \frac{(B)}{(B)} a_t$$

$$= (\quad\quad {}_{j}B^{j})a_t \quad {}_{j=0}$$

Recall?: The ratio of two polynomials is a (possibly infinite order) polynomial.

See Woodward et al. (2017) for more details concerning calculating ψ-weights.

DataScience@SMU

Calculating Psi Weights for ARMA Models

Example: $(1 - 1.2B + .6B^2)\, X_t = (1 - .5B)a_t$

can be written $X_t = \dfrac{1 \quad .5B}{1 \quad 1.2B + .6B^2} a_t$

Dividing polynomials (remember?)

Light board

$$\frac{1 \quad .5B}{1 \quad 1.2B + .6B^2}$$

$$= 1 - 1.2B + .6B^2 \overline{\big)\; 1 - .5B}$$

$$\begin{array}{r}
1 + .7B + .24B^2 - .132B^3 - \cdots \\
\hline
1 - .5B \\
\underline{1 - 1.2B + .6B^2} \\
.7B - .6B^2 \\
\underline{.7B - .84B^2 + .42B^3} \\
.24B^2 - .42B^3 \\
\underline{.24B^2 - .288B^3 + .144B^4} \\
-.132B^3 - .144B^4
\end{array}$$

Example: $(1 - 1.2B + .6B^2)\, X_t = (1 - .5B)a_t$

can be written $X_t = \dfrac{1 \quad .5B}{1 \quad 1.2B + .6B^2} a_t$

So, in this case, $\psi_0 = 1,\ \psi_1 = .7,\ \psi_2 = .24,\ \psi_3 = -.132,\ \ldots$

Light board

DataScience@SMU

tswge and Psi Weights

In practice: We will find ψ-weights using tswge function psi.weights.wge.

Return to example: $(1 - 1.2B + .6B^2)\, X_t = (1 - .5B)a_t$.

We calculated the ψ-weights to be $\psi_0 = 1,\ \psi_1 = .7,\ \psi_2 = .24,\ \psi_3 = -.132, \ldots$

Using tswge:

```
psi.weights.wge(phi=c(1.2,-.6),theta=.5,lag.max=5)
```

gives the output (for $\psi_1, \ldots, \psi_5$)

```
[1]  0.70000  0.24000 –0.13200 –0.30240 –0.28368
```

which agree with the above calculations.

Note: ψ_0 is always equal to 1.

DataScience@SMU

Probability Limits for ARMA Forecasts

Probability Limits for ARMA Forecasts

$$\frac{e_{t_0}(\ell)-0}{\sigma_a\left\{\sum_{j=0}^{\ell-1}\psi_j^2\right\}^{1/2}} \sim \text{N}(0,1)$$

Light board

i.e. $$\frac{X_{t_0+\ell}-\hat{X}_{t_0}(\ell)}{\sigma_a\left\{\sum_{j=0}^{\ell-1}\psi_j^2\right\}^{1/2}} \sim \text{N}(0,1)$$

$$\text{Prob}\left[\hat{X}_{t_0}(\ell)-z_{\alpha/2}\sigma_a\left\{\sum_{j=0}^{\ell-1}\psi_j^2\right\}^{1/2} \le X_{t_0+\ell} \le \hat{X}_{t_0}(\ell)+z_{\alpha/2}\sigma_a\left\{\sum_{j=0}^{\ell-1}\psi_j^2\right\}^{1/2}\right]=.95$$

DataScience@SMU

95% Probability Limits for Forecasts

$$\hat{X}_{t_0}(\ell) \pm z_{\alpha/2}\,\hat{\sigma}_a \left\{ \sum_{j=0}^{\ell-1} \psi_j^2 \right\}^{1/2}$$

Notes:

(1) 95% applies to individual forecasts

(2) The residuals, a_t, are unknown and must be estimated.

- tswge uses "backcasting" to find the $\hat{a}_t$'s.
- $\hat{\sigma}_a^2$ is the sample variance of the $\hat{a}_t$'s.
- See Woodward et al. (2017) Chapters 6 and 7.

Probability Limits for ARMA Forecasts with an Example

ARMA(2,1) example: $(1 \quad 1.2B + .6B^2)(X_t \quad 50) = (1 \quad .5B)a_t$

$$\boxed{\hat{X}_{t_0}(\ell) \pm 1.96 \ \hat{\sigma}_a \left\{ \sum_{j=0}^{\ell-1} \psi_j^2 \right\}^{1/2}}$$

Estimate $_a$ using $\hat{\ }_a = .8659$

-weights: $_0 = 1, \quad _1 = .7,$

$_2 = .24, \quad _3 = \quad .132$

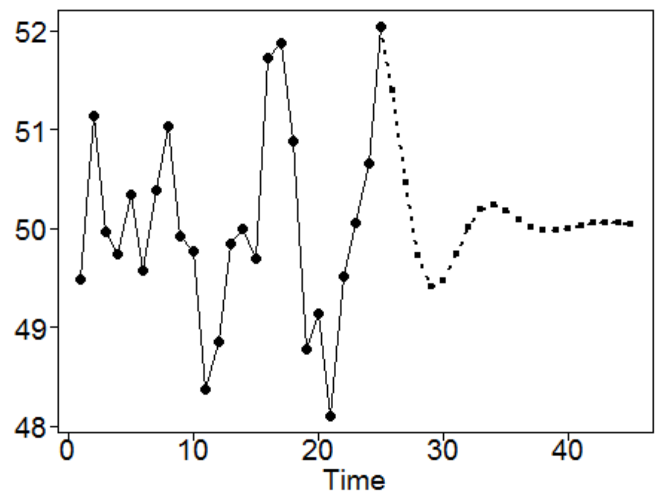
$\ell = 1 \quad 1.96(.8659)(1) = 1.70$

$\ell = 2 \quad 1.96(.8659)\sqrt{1 + .7^2} = 2.07$

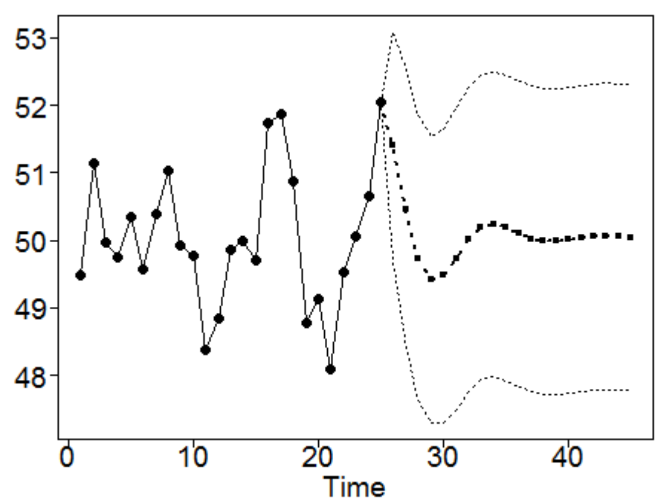
$\ell = 3 \quad 1.96(.8659)\sqrt{1 + .7^2 + .24^2} = 2.11$

ℓ	Forecast	Half-width
1	51.40	1.70
2	50.46	2.07
3	49.73	2.11
4	49.42	2.12

ARMA(2,1) example: $(1 \quad 1.2B + .6B^2)(X_t \quad 50) = (1 \quad .5B)a_t$



Forecasts with 95% probability limits



DataScience@SMU

tswge and Probability Limits (aka Forecast Limits)

tswge demo: showing forecast limits

To forecast from a stationary ARMA model and show 95% probability limits, use

```
fore.arma.wge(x,phi,theta,n.ahead,plot=TRUE,
limits=TRUE)
```

Example: Forecasts using the AR(1) model $(1 \quad .8B)(X_t \quad 25) = a_t$

```
data(fig6.1nf)
fore.arma.wge(fig6.1nf,phi=.8,n.ahead=20,
plot=TRUE,limits=FALSE)

fore.arma.wge(fig6.1nf,phi=.8,n.ahead=20,
plot=TRUE,limits=TRUE)
```

tswge demo

ARMA(2,1) example: $(1 \quad 1.2B + .6B^2)(X_t \quad 50) = (1 \quad .5B)a_t$

```
data(fig6.2nf)
fore.arma.wge(fig6.2nf,phi=c(1.2,-.6), theta=.5,
n.ahead=20,plot=TRUE,limits=FALSE)

fore.arma.wge(fig6.2nf,phi=c(1.2,-.6), theta=.5,
n.ahead=20,plot=TRUE,limits=TRUE)
```

DataScience@SMU

Checking Forecasts

ASE

Checking Forecasts

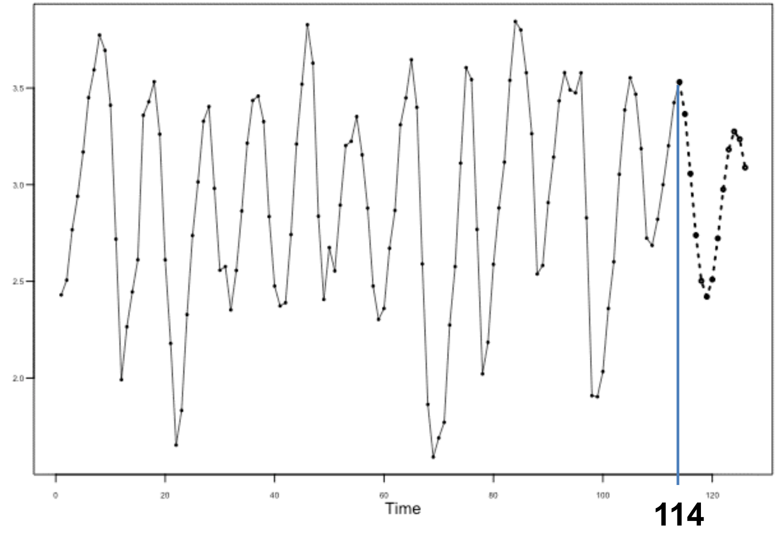
If *n* is the length of your time series realization, then sometimes it is useful to set $t_0 = n - k$ and "forecast the last k values in the realization."

- In this way, you can "check" your forecasts with the actual values.
- tswge function allows you to do this using the "lastn" option.
  - If lastn=F, then you forecast past the end of the realization as we have in the previous examples (lastn=F is the default option).
  - If lastn=T, then you forecast the last n.ahead values in the series.

Checking Forecasts

Again, consider the log lynx dataset. Previously, we have looked at an AR(4) model for these data:

$$(1 - 1.3B + 0.7B^2 - 0.1B^3 + 0.2B^4)(X_t - 2.9) = a_t$$



f = fore.arma.wge(llynx,phi = c(1.3, -0.7, 0.1, -0.2), n.ahead = 12, limits = FALSE)

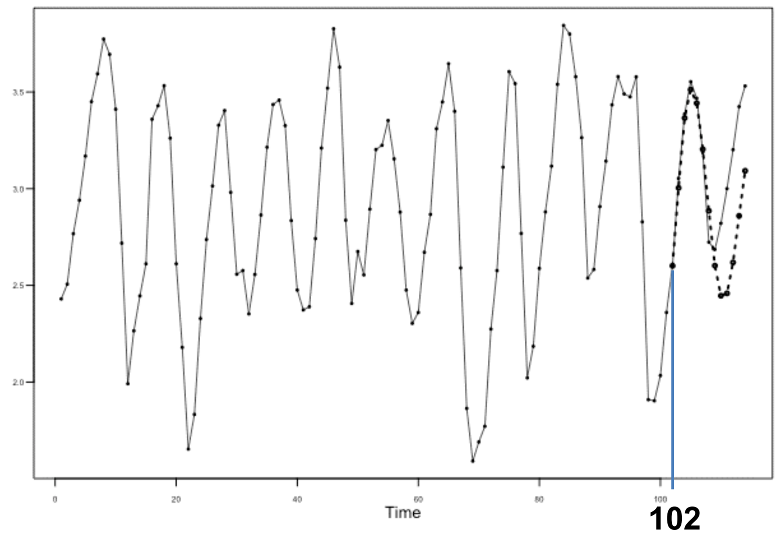
But how good are these forecasts?

We can’t know because we don’t know the actual values.

Checking Forecasts

Again, consider the log lynx dataset. Previously, we have looked at an AR(4) model for these data:

$$(1 - 1.3B + 0.7B^2 - 0.1B^3 + 0.2B^4)(X_t - 2.9) = a_t$$




```
f = fore.arma.wge(llynx,phi = c(1.3, -0.7, 0.1, -0.2), n.ahead = 12, lastn = TRUE, limits = FALSE)
```

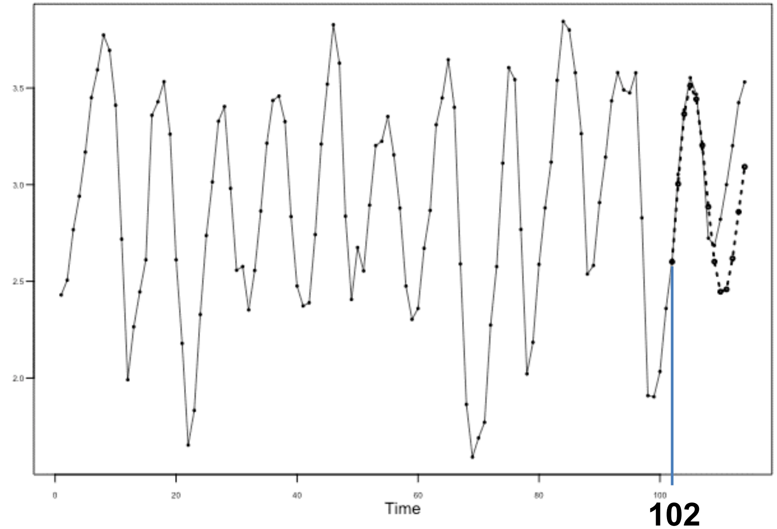
We can forecast the last n (12 in this case) values of the series so that we can evaluate the fit.

We will use the average square error (ASE): $ASE = \dfrac{\sum(\hat{X}_i - X_i)^2}{n}$

Checking Forecasts

Again, consider the log lynx dataset. Previously, we have looked at an AR(4) model for these data:

$$(1 - 1.3B + 0.7B^2 - 0.1B^3 + 0.2B^4)(X_t - 2.9) = a_t$$



f = fore.arma.wge(llynx,phi = c(1.3, -0.7, 0.1, -0.2), n.ahead = 12, lastn = TRUE, limits = FALSE)

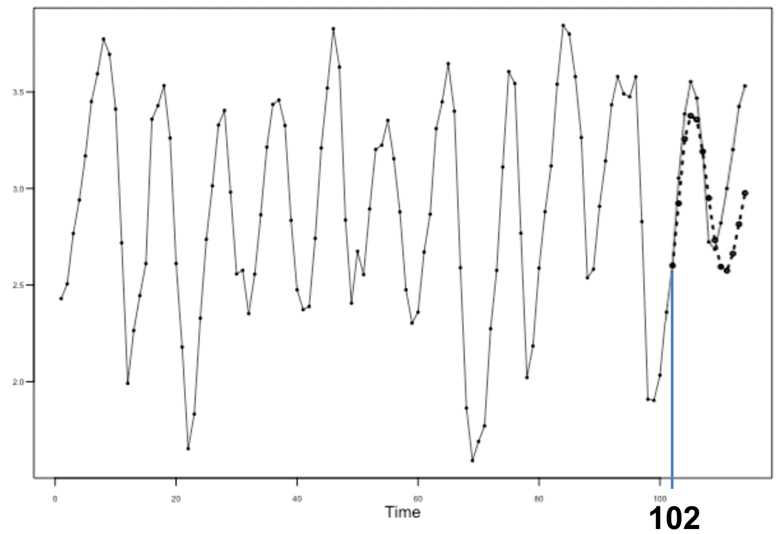
$$ASE = \frac{\sum(\hat{X}_i - X_i)^2}{n} = \frac{\sum_{i=1}^{12}(\hat{X}_i - X_i)^2}{12} \rightarrow$$

```
> ASE = mean((f$f-llynx[103:114])^2)
> ASE
[1] 0.1102716
```

Checking Forecasts

Next, consider the ARMA(4,1) model we looked at for this data set. Let's again calculate the ASE for the last 12 observations.

$$1 - 0.7B - 0.1B^2 + 0.2B^3 + 0.3B^4)(X_t - 2.9) = (1 + .6B)a_t$$



f = fore.arma.wge(llynx, phi = c(0.7, 0.1, - 0.2, - 0.3), theta = -.6, n.ahead = 12, lastn = TRUE, limits = FALSE)

$$ASE = \frac{\sum(\hat{X}_i - X_i)^2}{n} \rightarrow$$

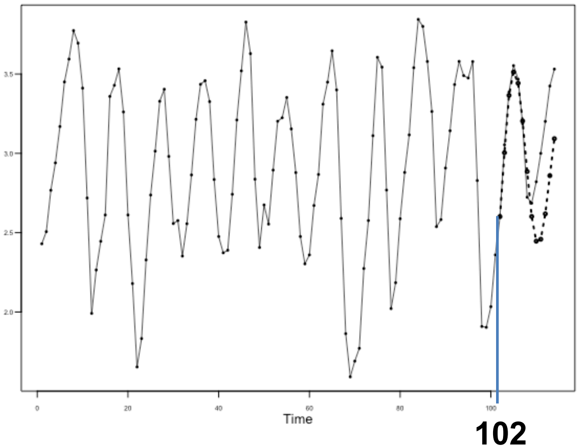
```
> ASE = mean((f$f-llynx[103:114])^2)
> ASE
[1] 0.1109845
```

Checking Forecasts

Next, consider the ARMA(4,1) model we looked at for this data set. Let's again calculate the ASE for the last 12 observations.

AR(4)

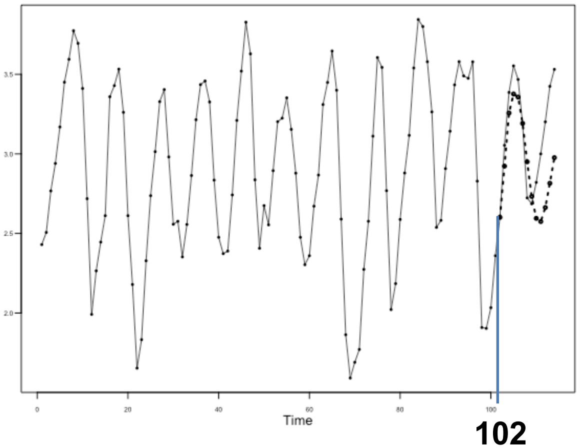
$$(1 - 1.3B + 0.7B^2 - 0.1B^3 + 0.2B^4)(X_t - 2.9) = a_t$$




```
> ASE = mean((f$f-llynx[103:114])^2)
> ASE
[1] 0.1102716
```

ARMA(4,1)

$$1 - 0.7B - 0.1B^2 + 0.2B^3 + 0.3B^4)(X_t - 2.9) = (1 + .6B)a_t$$




```
> ASE = mean((f$f-llynx[103:114])^2)
> ASE
[1] 0.1109845
```

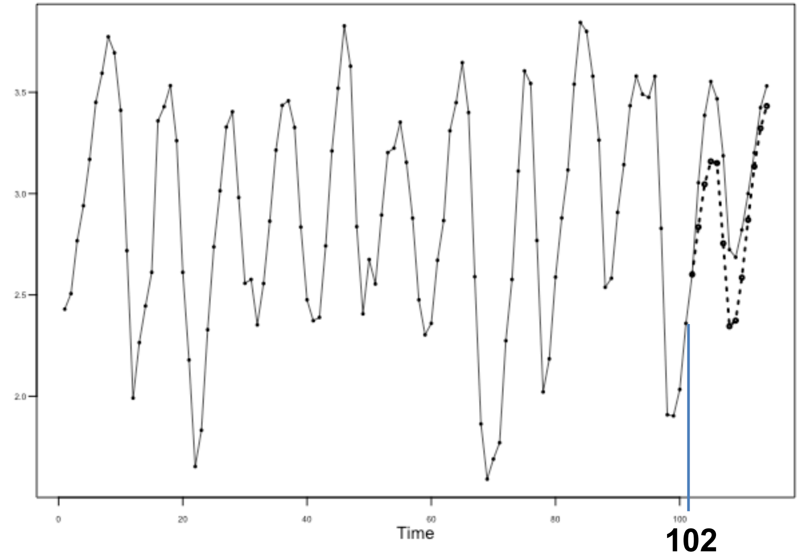
The ASE for the AR(4) model is smaller; therefore, in terms of ASE, the evidence suggests the AR(4) model fits better that the AR(4,1).

Example: Canadian Lynx Data

A classic model fit to the log lynx data is an AR(11) (Tong, 1977).

```
data(llynx)
f = fore.arma.wge(llynx,phi=c(1.17, -0.54, 0.27, -0.31, 0.15, -0.15,
0.06, -0.03,0.13, 0.20, - 0.34),n.ahead=12,limits=FALSE, lastn = TRUE)
```

Forecasting last 12 values

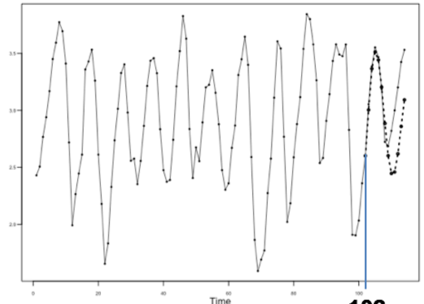



```
mean((f$f-llynx[114-12+1])^2
[1] 0.07865787
```

Example: Canadian Lynx Data

AR(4)

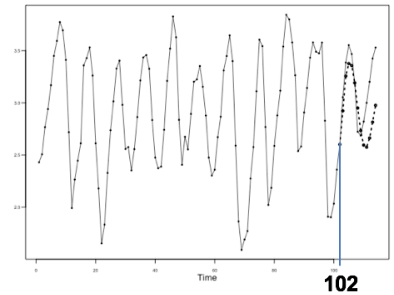
$(1 - 1.3B + 0.7B^2 - 0.1B^3 + 0.2B^4)(X_t - 2.9) = a_t$




```
> ASE = mean((f$f-llynx[103:114])^2)
> ASE
[1] 0.1102716
```

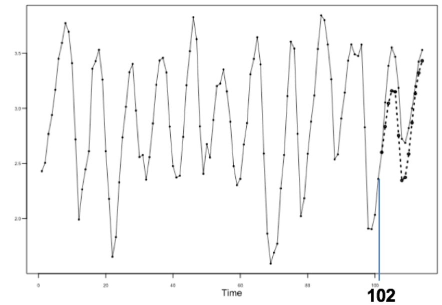
ARMA(4,1)

$1 - 0.7B - 0.1B^2 + 0.2B^3 + 0.3B^4)(X_t - 2.9) = (1 + .6B)a_t$




```
> ASE = mean((f$f-llynx[103:114])^2)
> ASE
[1] 0.1109845
```

Forecasting last 12 values from AR(11)




```
mean((f$f-llynx[114-12+1])^2
[1] 0.07865787
```

DataScience@SMU

tswge

Checking Forecasts

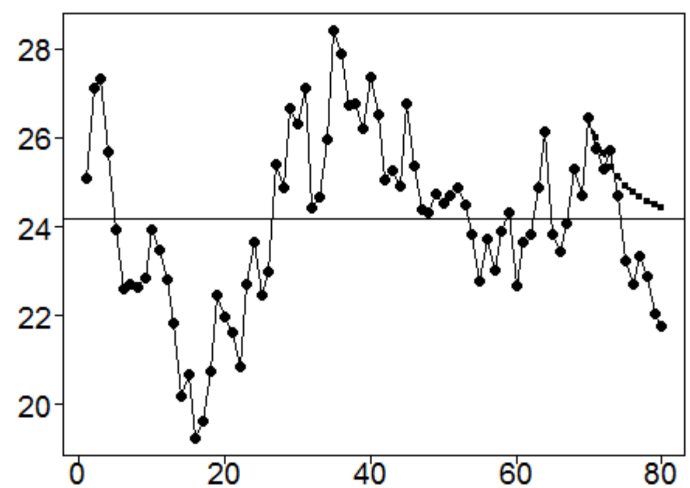
tswge demo: “forecasting” last n.ahead values in your realization

To forecast the last n.ahead values in your data set, use

`fore.arma.wge(x,phi,theta,n.ahead,lastn=T, plot=T,limits=F)`

Example: Forecasts using the AR(1) model $(1 - .8B)(X_t - 25) = a_t$

```
data(fig6.1nf)
fore.arma.wge(fig6.1nf,phi=.8,n.ahead=10,lastn=TRUE,
plot=TRUE,limits=FALSE)
```



tswge demo: “forecasting” last n.ahead values in your realization

To forecast the last n.ahead values in your data set, use

```
fore.arma.wge(x,phi,theta,n.ahead,lastn=TRUE)
```

ARMA(2,1) example: $(1 \quad 1.2B+.6B^2)(X_t \quad 50)=(1 \quad .5B)a_t$

The ARMA(2,1) data set in fig6.2nf only has 25 data values.

- In order to illustrate the “lastn” feature, we generate a longer realization (n=100) from this model.
- Notice in the code below that we used sn=8 so that this particular realization can be regenerated.
- On the following slide, we illustrate forecasting the last 5, 10, and 20 data values.

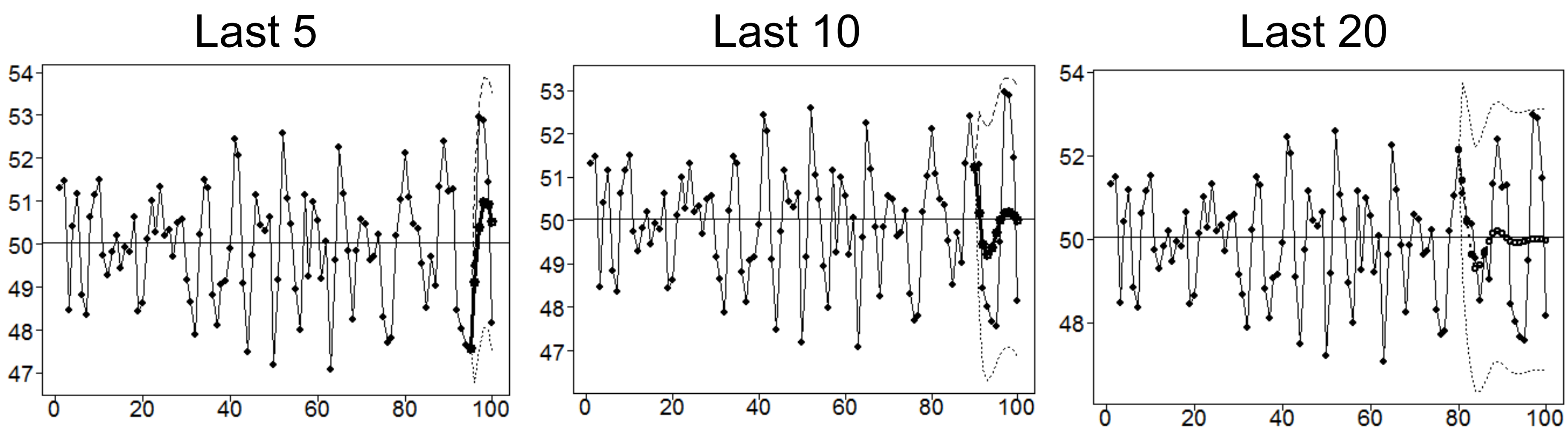
tswge demo: “forecasting” last n.ahead values

ARMA(2,1) example: $(1 \quad 1.2B+.6B^2)(X_t \quad 50)=(1 \quad .5B)a_t$

```
x21=gen.arma.wge(n=100,phi=c(1.2,-.6),theta=.5,sn=8)
x21=x21+50
fore.arma.wge(x21,phi=c(1.2,-.6),theta=.5,n.ahead=5, lastn=TRUE)
```

We can recreate this realization.

We also use 10 and 20.



tswge demo: “forecasting” last n.ahead values

Notes:

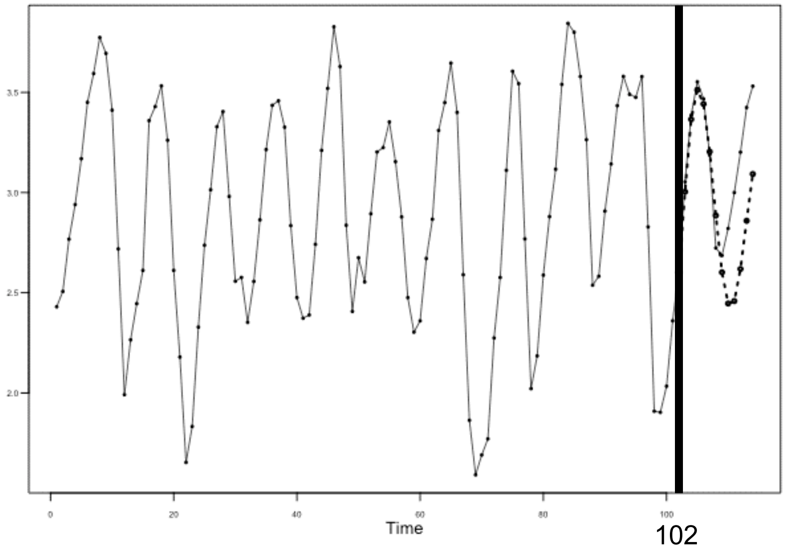
- Short-term forecasts tended to be pretty good.
- For longer steps ahead, the forecasts tended to the mean, 50, shown by the horizontal line.
- In all cases, the actual values were within the 95% probability limits.

Example: Canadian Lynx Data

A classic model fit to the log lynx data is an AR(4) (Tong, 1977).

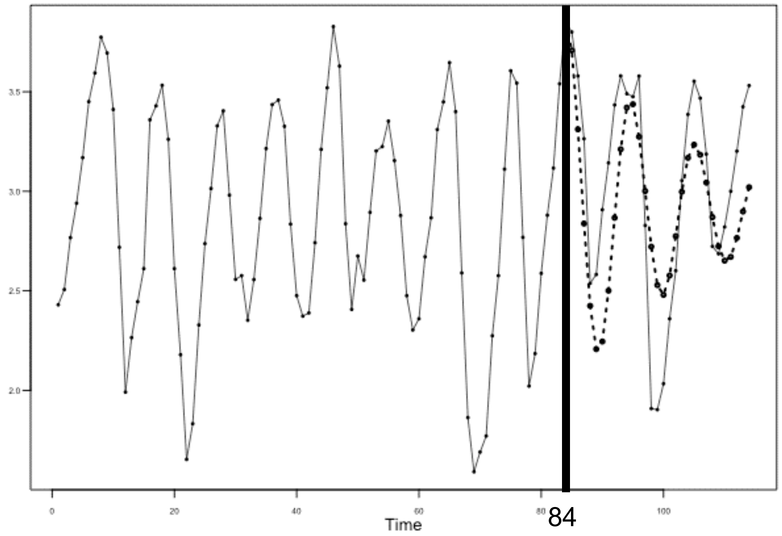
```
data(llynx)
f = fore.arma.wge(llynx,phi=c(1.3, -0.7,0.1,0.2), n.ahead=30,
limits=FALSE, lastn = TRUE)
```

Forecasting last 12 values




```
> mean((llynx[(114 - 12 + 1):114] - f$f)^2)
[1] 0.1102716
```

Forecasting last 30 values



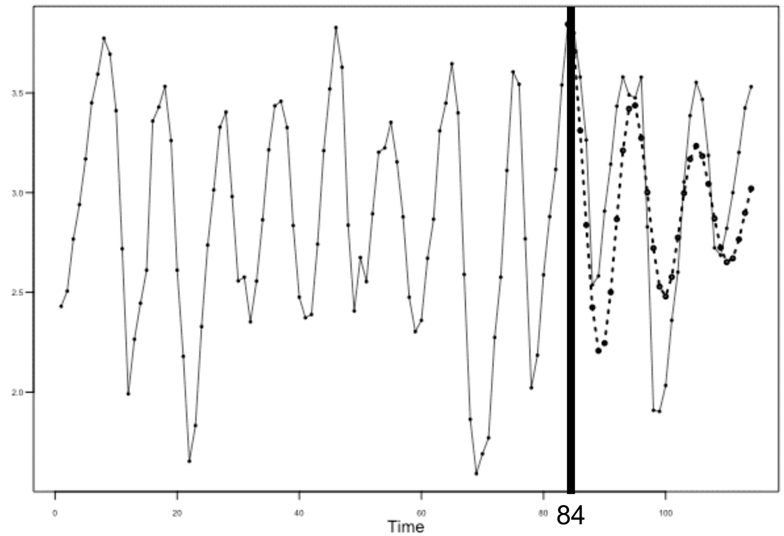

```
> mean((llynx[(114 - 30 + 1):114] - f$f)^2)
[1] 0.1440614
```

Example: Canadian Lynx Data

Another classic model fit to the log lynx data is an AR(11) (Tong, 1977).

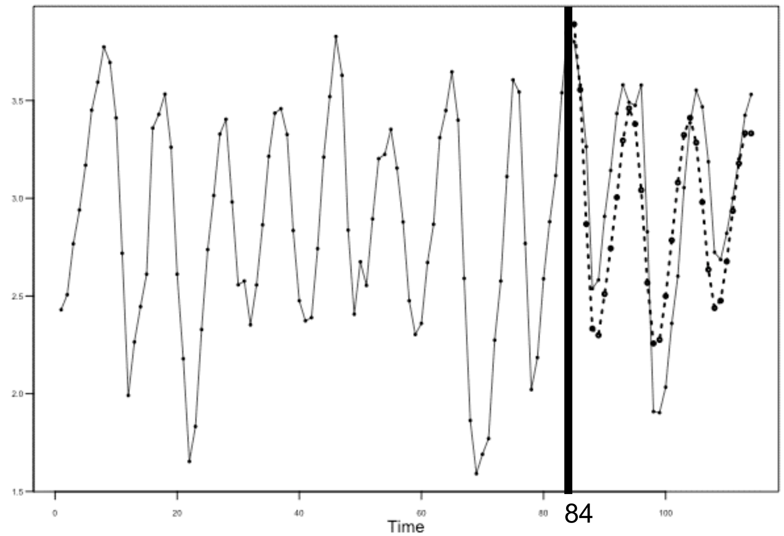
```
data(llynx)
f = fore.arma.wge(llynx,phi=c(1.17, -0.54, 0.27, -0.31, 0.15, -0.15,
0.06, -0.03,0.13, 0.20, -0.34),n.ahead=30,limits=FALSE, lastn = TRUE)
```

Forecasting last 30 values from AR(4)




```
> mean((llynx[(114 - 30 + 1):114] - f$f)^2)
[1] 0.1440614
```

Forecasting last 30 values from AR(11)




```
> mean((llynx[(114 - 30 + 1):114] - f$f)^2)
[1] 0.1003434
```

DataScience@SMU

Introduction

Forecasting with ARIMA

Forecasting with ARIMA

$$(B)(1 \quad B)^d (X_t \qquad) = \ (B)a_t$$

Comments:

- forecasts from these models are obtained using a method analogous to that used for ARMA forecasts
- we will not calculate these forecasts by hand but will typically use tswge function fore.aruma.wge
- forecast limits for these nonstationary models become unbounded as ℓ increases

DataScience@SMU

Example

Forecasting with ARIMA(0,1,0)

Light board

Example: $(1 \quad B)(X_t \qquad) = a_t$

Recall: For stationary AR(1)—for example, $(1-.9B)(X_t - \mu) = a_t$

$$\hat{X}_{t_0}(\ell) = .9\hat{X}_{t_0}(\ell \quad 1) + \bar{X}(1 \quad .9)$$

Our case: $(1 \quad B)(X_t \qquad) = a_t$ i.e. $\varphi_1 = 1$ instead of $.9$

$$\hat{X}_{t_0}(\ell) = \hat{X}_{t_0}(\ell \quad 1) + \bar{X}(1 \quad 1)$$ ($\bar{X}$ plays no role in the forecasts)

$$\hat{X}_{t_0}(\ell) = \hat{X}_{t_0}(\ell \quad 1)$$

$$\hat{X}_{t_0}(1) = \hat{X}_{t_0}(0) = X_{t_0}$$

$$\hat{X}_{t_0}(2) = \hat{X}_{t_0}(1) = X_{t_0}$$

$$\vdots$$

Continues to forecast the last data value

- Boring
- But makes sense

DataScience@SMU

Example

Forecasting with ARIMA(0,1,0)

Example: $(1 \quad B)(X_t \qquad) = a_t$

Recall: For stationary AR(1)—for example, $(1-.9B)(X_t-\mu) = a_t$

$$\hat{X}_{t_0}(\ell) = .9\hat{X}_{t_0}(\ell \quad 1) + \bar{X}(1 \quad .9)$$

Our case: $(1 \quad B)(X_t \qquad) = a_t$ i.e. $\varphi_1 = 1$ instead of $.9$

$\hat{X}_{t_0}(\ell) = \hat{X}_{t_0}(\ell \quad 1) + \bar{X}(1 \quad 1)$ ($\bar{X}$ plays no role in the forecasts)

$\hat{X}_{t_0}(\ell) = \hat{X}_{t_0}(\ell \quad 1)$

$\hat{X}_{t_0}(1) = \hat{X}_{t_0}(0) = X_{t_0}$

$\hat{X}_{t_0}(2) = \hat{X}_{t_0}(1) = X_{t_0}$

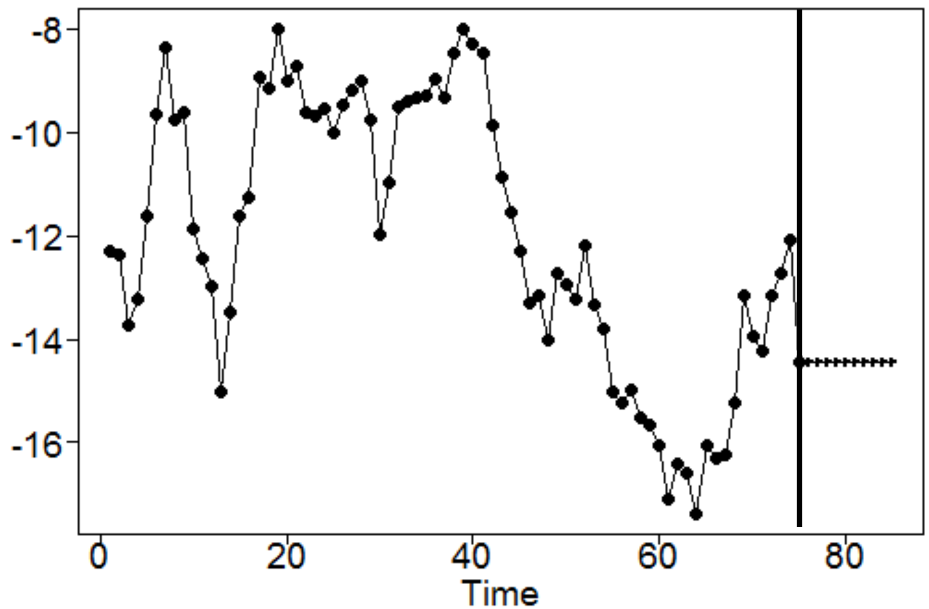
$\vdots$

Continues to forecast the last data value

- Boring
- But makes sense

Example: $(1 \quad B)(X_t \qquad) = a_t$

```
xd1=gen.aruma.wge(n=75,d=1,sn=74)
fore.aruma.wge(xd1,d=1,n.ahead=5,limits=F)
```

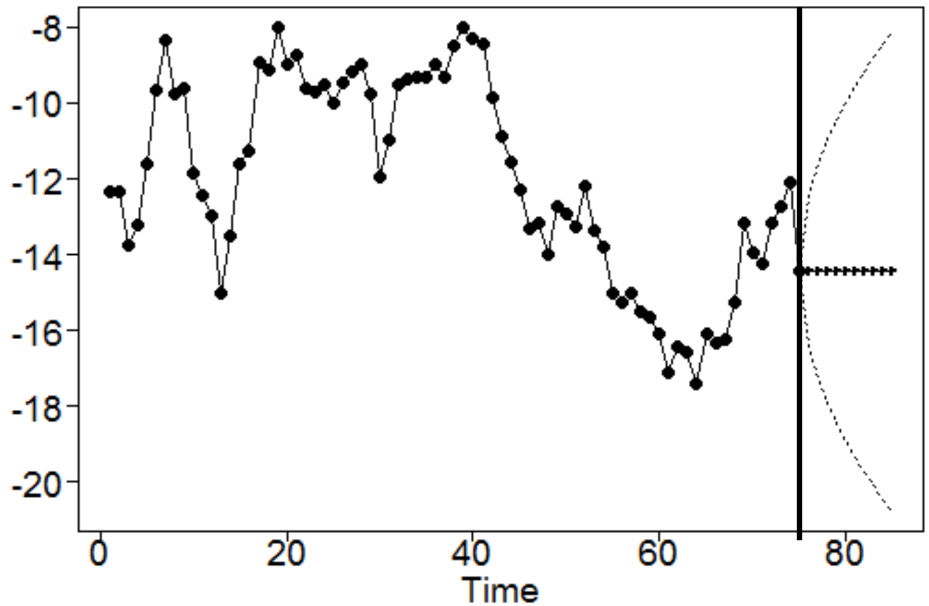


Based on the realization, it’s unclear whether one should forecast the data values to go up or down.

- Forecasting the last value seems reasonable.

Example: $(1 \quad B)(X_t \qquad) = a_t$

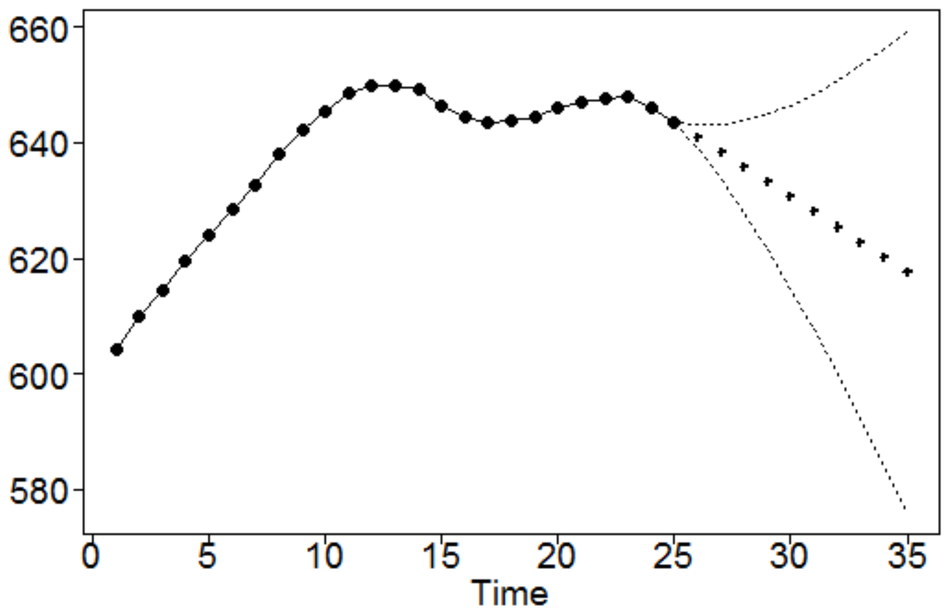
xd1=gen.aruma.wge(n=75,d=1,sn=74)
fore.aruma.wge(xd1,d=1,n.ahead=5,limits=T)



Selecting 95% Probability Limits

- Shows the uncertainty associated with the forecasts
    - Could go up or down
- Limits seem to be increasing without bound

Example: $(1 \quad B)^2(X_t \qquad) = a_t$



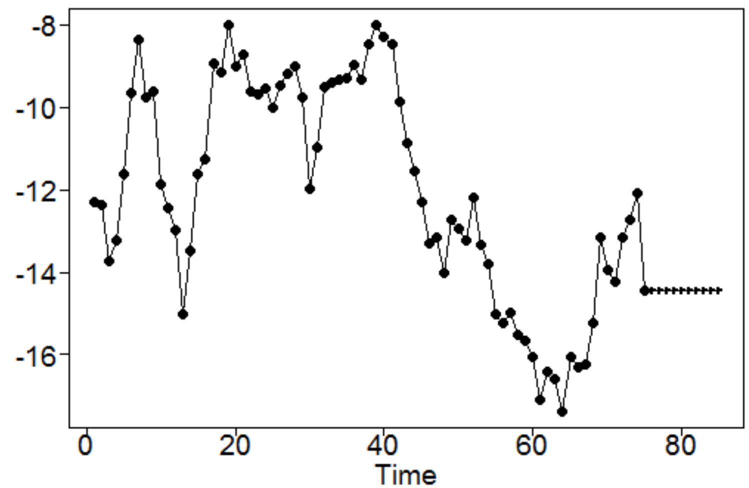
Note: Forecasts follow a line determined by the last 2 data values.

- That is, the most recent trend is predicted to continue.

Key Concept!

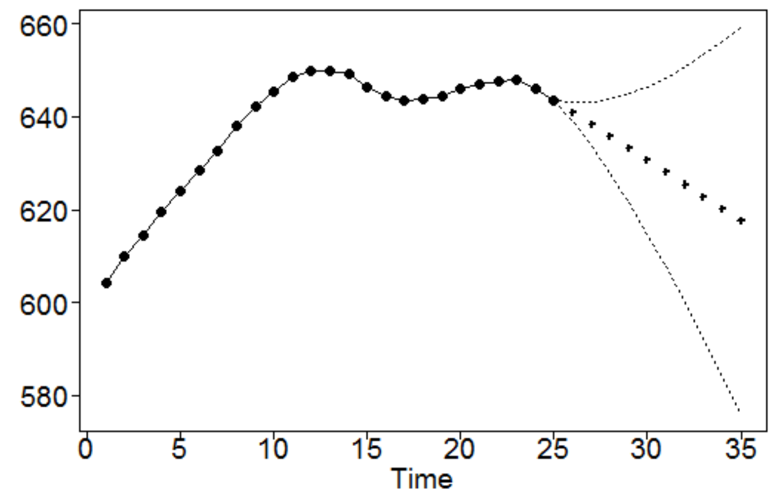
$$(1 \quad B)(X_t \qquad) = a_t$$

Forecasts the the last value



$$(1 \quad B)^2(X_t \qquad) = a_t$$

Forecasts the trend determined by the last two observations



DataScience@SMU

tswge

Generating and Forecasting Simple ARIMA Models

```
fore.aruma.wge(x,phi,theta,d,s,lambda,n.ahead,
lastn=FALSE,plot=TRUE,limits=TRUE)
```

Example: forecasts using the ARIMA(0,1,0) model

$$(1-B)\,X_t = a_t$$

tswge demo

```
x=gen.aruma.wge(n=50,phi=.8,d=1,sn=15)
fore.aruma.wge(x,d=1,n.ahead=20 , limits = FALSE)
```

Example: forecasts using the ARIMA(1,1,0) model

$$(1-.8B)(1-B)\,X_t = a_t$$

```
x=gen.aruma.wge(n=50,phi=.8,d=1,sn=15)
fore.aruma.wge(x,phi=.8,d=1,n.ahead=20 , limits = FALSE)
```

Example: forecasts using the ARIMA(0,2,0) model

$$(1-B)^2\,X_t = a_t$$

```
x=gen.aruma.wge(n=50,phi=.8,d=1,sn=15)
fore.aruma.wge(x,d=2,n.ahead=20, limits = FALSE)
```

DataScience@SMU

Introduction

Forecasting with Seasonal ARIMA Models

Forecasting with Seasonal Models

$$\left(1 \quad B^{s}\right)\left(X_{t} \qquad\right) = a_{t}$$

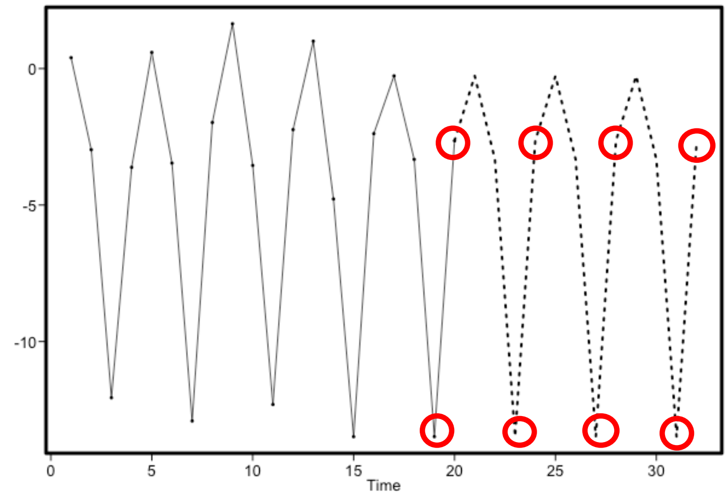
Example: $\left(1 \quad B^{4}\right)\left(X_{t} \qquad\right) = a_{t}$

$$\hat{X}_{t_0}(\ell) = \hat{X}_{t_0}(\ell - 4) + \bar{X}(1 - 0 - 0 - 0 - 1) \qquad \hat{X}_{t_0}(\ell) = \sum_{j=1}^{p} \varphi_j \hat{X}_{t_0}(\ell - j) - \sum_{j=\ell}^{q} \theta_j \hat{a}_{t_0+\ell-j} + \bar{X}\left[1 - \sum_{i=1}^{p} \varphi_i\right]$$

$$\hat{X}_{t_0}(\ell) = \hat{X}_{t_0}(\ell - 4)$$

Note:

Given quarterly data, this very simple model forecasts the current quarter to be the value observed for this quarter last year.



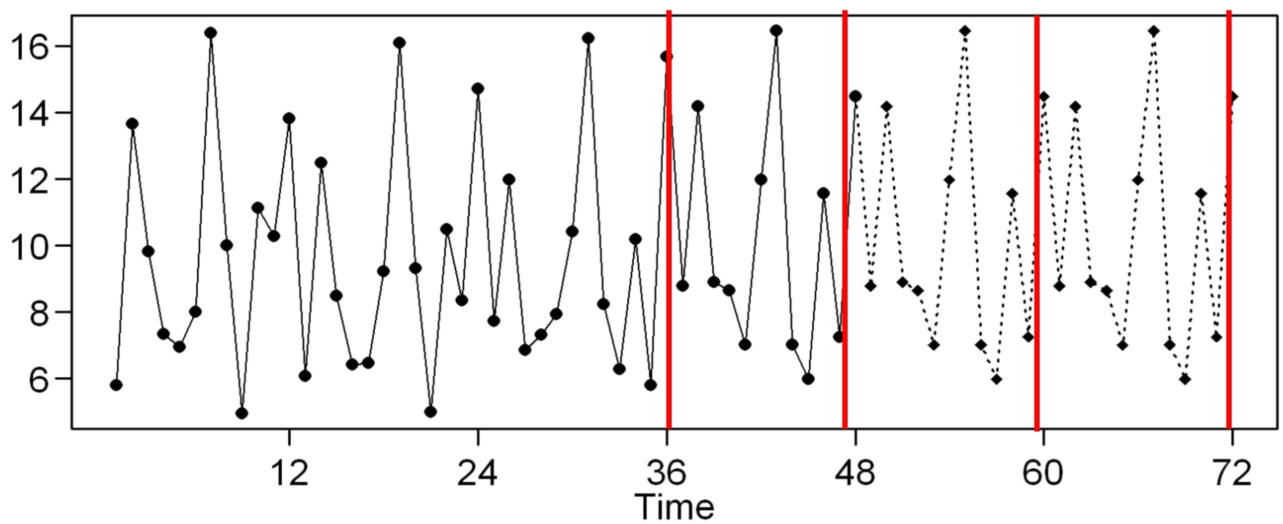
Example:

$$\left(1 \quad B^{s}\right)\left(X_{t} \qquad\right) = a_{t}$$

Forecasts: $\hat{X}_{t_0}(\ell) = \hat{X}_{t_0}(\ell - s)$

- That is, forecasts are an exact replica of the last s values.

Realization and forecasts for $(1 \quad B^{12})(X_t \qquad) = a_t$
as of time $t_0 = 36$



DataScience@SMU

tswge

Generating and Forecasting Simple Seasonal Models

DataScience@SMU

tswge demo

`fore.aruma.wge(x,phi,theta,d,s,n.ahead,`
`lastn=FALSE,plot=TRUE,limits=TRUE)`

Forecasts from the pure seasonal model $(1-B^4)X_t = a_t$

```
x=gen.aruma.wge(n=20,s=4, sn = 6)
fore.aruma.wge(x,s=4,n.ahead=8,lastn=FALSE,plot=TRUE,
limits=FALSE)
#
x=gen.aruma.wge(n=20,s=4, sn = 6)
fore.aruma.wge(x,s=4,n.ahead=8,lastn=TRUE,plot=TRUE,
limits=FALSE)
```

Example: forecasts using seasonal model $(1-.8B)(1-B^4)X_t = a_t$

```
x=gen.aruma.wge(n=20,phi=.8,s=4,sn = 6)
fore.aruma.wge(x,phi=.8,s=4,n.ahead=8,limits=FALSE)
```

Also: `lastn=TRUE`

DataScience@SMU

A More General Seasonal Model/ the "Airline Model"

Factor Table: s=12

Factor table for $(1 - B^{12})$

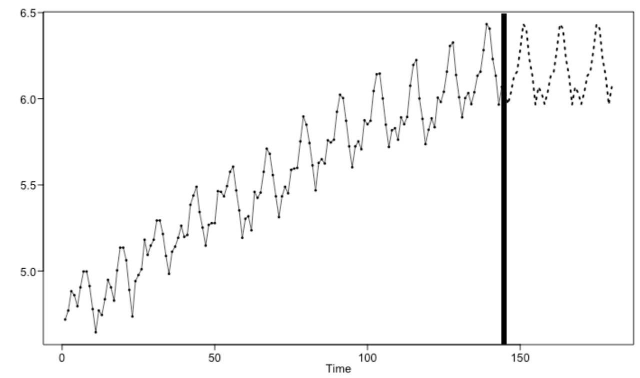
Factor	Abs Recip	f	Root(s)
$1 - B$	1	0	1
$1 - \sqrt{3}B + B^2$	1	.083	$.866 \pm .5i$
$1 - B + B^2$	1	.167	$.5 \pm .866i$
$1 + B^2$	1	.25	$\pm\, i$
$1 + B + B^2$	1	.333	$-.5 \pm .866i$
$1 + \sqrt{3}B + B^2$	1	.417	$-.866 \pm .5i$
$1 + B$	1	.5	-1

Note: $(1 - B^{12})$ has a factor of $(1 - B)$ so that the product $(1 - B)(1 - B^{12})$ has **two** factors of $(1 - B)$.

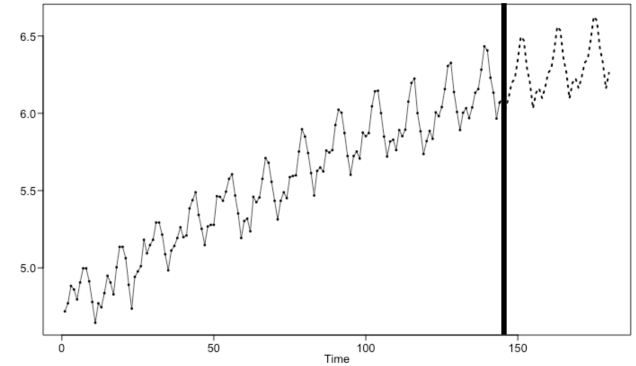
Forecasts Using the Airline Model

- Using this model, the data for the last year are forecast into the future.
- The trend is not forecast to continue.
- In our previous examples:
  - $(1 - B)\, X_t = a_t$ simply forecasts the last data value into the future
  - $(1 - B)^2\, X_t = a_t$ forecasts the trend to continue

$$(1 - B^{12})(X_t - \mu) = a_t$$



$$(1 - B)(1 - B^{12})(X_t - \mu) = a_t$$



More general seasonal model

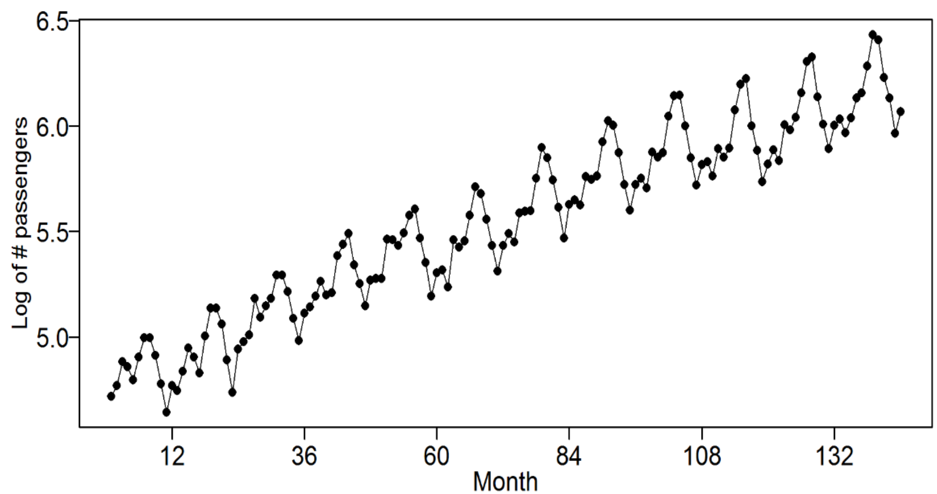
$$(B)(1 \quad B)(1 \quad B^s)(X_t \qquad) = \quad (B)a_t$$

- Forecasts follow general pattern of last s points, but not an exact replica
- “Finer detail” incorporated through $\varphi(B)$ and $\theta(B)$

Example: Seasonal Model with Trend

$$(B)(1 \quad B)(1 \quad B^s)(X_t \qquad) = \quad (B)a_t$$

- It is called the airline model.
- It is useful for modeling and forecasting data similar to the (log) airline data shown below.
- Notice that a key feature of this data set is that there is a seasonal component and a trend



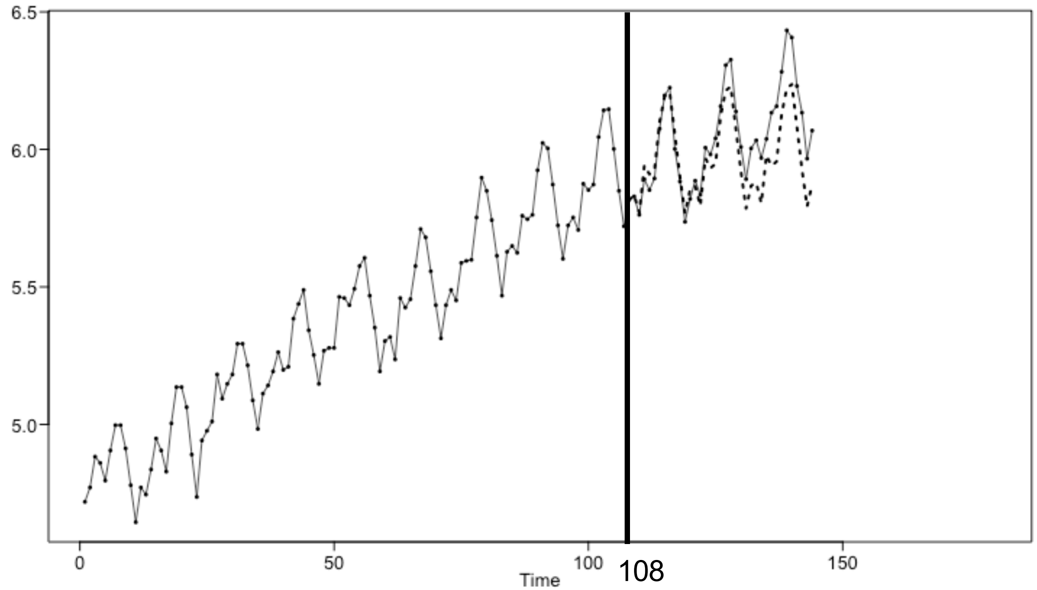
DataScience@SMU

Parzen, Box and Woodward / Gray Airline Models

$$\varphi(B)\,(1-B^{12})(X_t-\mu)=a_t$$

Where $\varphi(B)$ is the 13th-order AR operator given by Parzen et al. (1980).

```
data(airlog)
fore.aruma.wge(airlog, d = 0, s = 12, phi = c(.74,0,0,0,0,0,0,0,0,0,0,.38,-.2812),n.ahead = 36,lastn = TRUE, limits = FALSE)
```

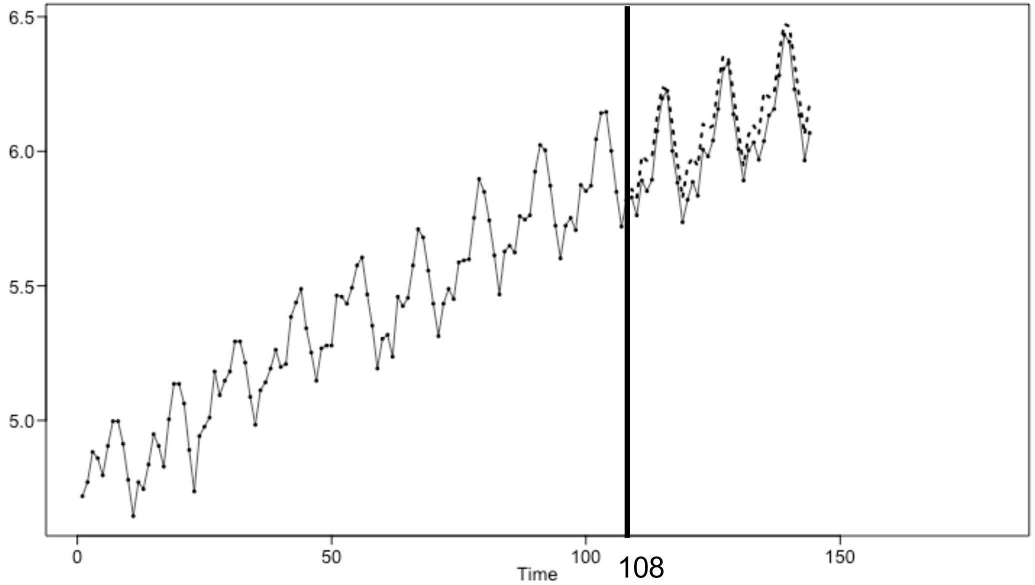


Seasonality predicted to continue… but not the trend

$$(1 - B)(1 - B^{12})(X_t - \mu) = \theta(B)a_t$$

Where $\theta(B)$ is the 13th-order MA operator given by Box et al. (2008).

```
data(airlog)
fore.aruma.wge(airlog, d = 0, s = 12, phi = c(.74,0,0,0,0,0,0,0,0,0,0,.38,-.2812),n.ahead = 36,lastn = TRUE, limits = FALSE)
```

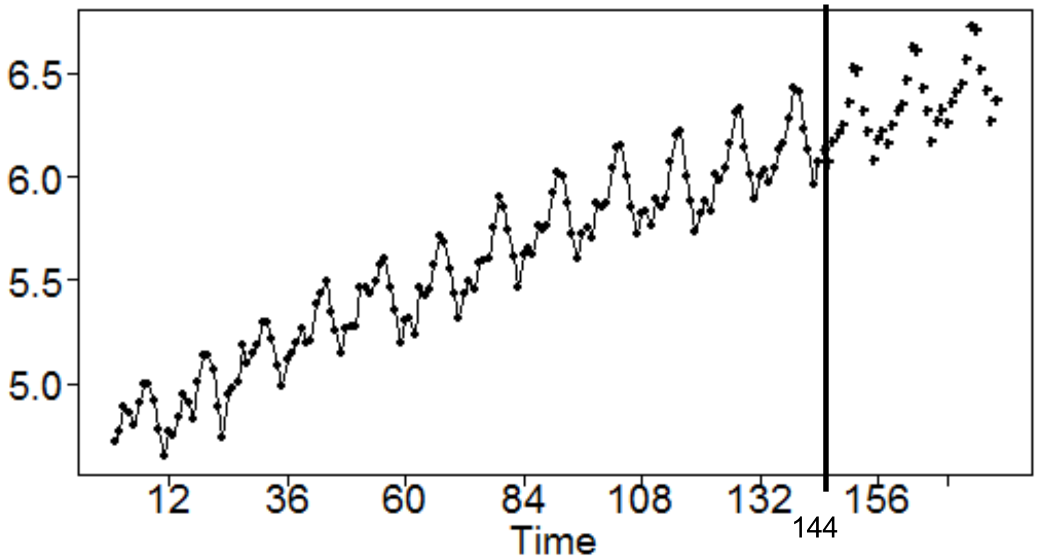


Seasonality predicted to continue… as well as the trend.

$$\varphi_1(B)(1 - B)(1 - B^{12})\, X_t = a_t$$

Where $\varphi_1(B)$ is the 12th-order operator given by Woodward et al. (2017).

```
data(airlog)
phi1=c(-.36,-.05,-.14,-.11,.04,.09,-.02, .02,.17,.03,-.10,-.38)
fore.aruma.wge(airlog,phi=phi1,d=1,s=12,n.ahead=36,plot=T,lastn=F,limits=F)
```

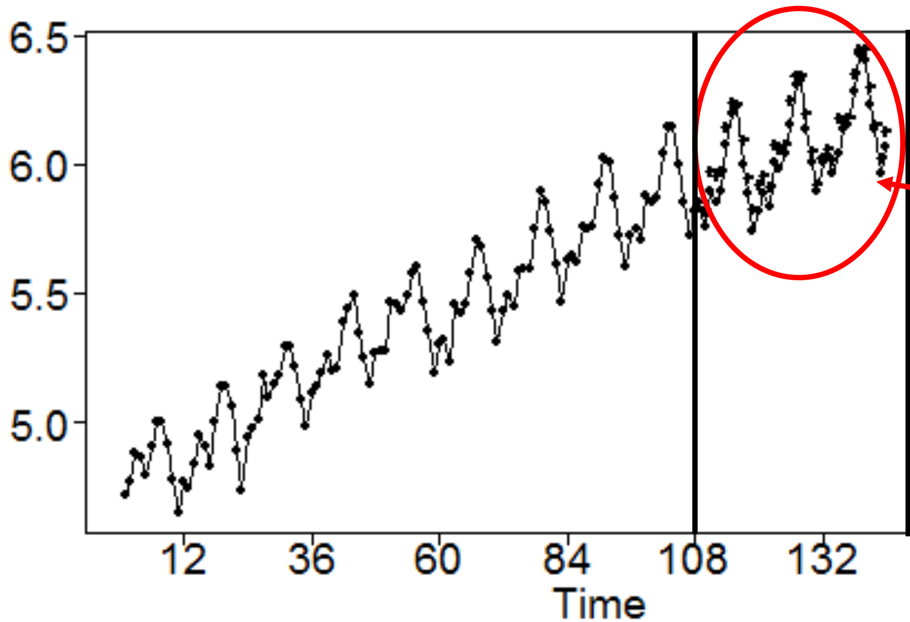


Seasonality **and** trend predicted to continue

$$\varphi_1(B)(1-B)(1-B^{12})\,X_t = a_t$$

Where $\varphi_1(B)$ is the 12th-order operator given by Woodward et al. (2017).

```
data(airlog)
phi1=c(-.36,-.05,-.14,-.11,.04,.09,-.02, .02,.17,.03,-.10,-.38)
fore.aruma.wge(airlog,phi=phi1,d=1,s=12,n.ahead=36,plot=T,lastn=T,limits=F)
```



Forecasts of last 36 months are so good that it's difficult to distinguish them from the actual values.

DataScience@SMU

Competition

Parzen, Box, Woodward / Gray — Airline Forecasts, AIC, and ASE

Airline Competition Among the Greats!

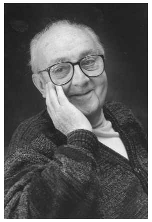
Previously, we looked at the structure of the model from three world-renowned statisticians.

```
> data(airlog)
> SA1 = artrans.wge(airlog,1) # take first differences of the data
> SA1_12 = artrans.wge(SA1,c(rep(0,11),1)) # take the 12th difference of the first difference (1-B)(1-B12)
> SA12 = artrans.wge(airlog,c(rep(0,11),1)) # take the 12th difference of the data (1- B12)
> Parzen = aic.wge(SA12, p = 13) #Phi(B)(1-B12)(Xt-mu) = at
> Box = aic.wge(SA1_12, q = 13)  #(1-B)(1-B12) (Xt-mu) = Theta(B)at
> Woodward = aic.wge(SA1_12, p = 12) # #Phi(B)(1-B)(1-B12) (Xt-mu) = at
> Parzen$value
[1] -6.465559
> Box$value
[1] -6.499275
> Woodward$value
[1] -6.423649
> # We will return to this competition when we use the model to forecast!
```

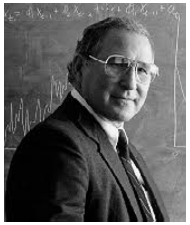
Airline Competition Among the Greats!

```
> Parzen$value
[1] -6.465559
> Box$value
[1] -6.499275
> Woodward$value
[1] -6.423649
```

$$(1 - .74B)(1 + .38B^{12})(1 - B^{12})(X_t - \mu) = a_t$$



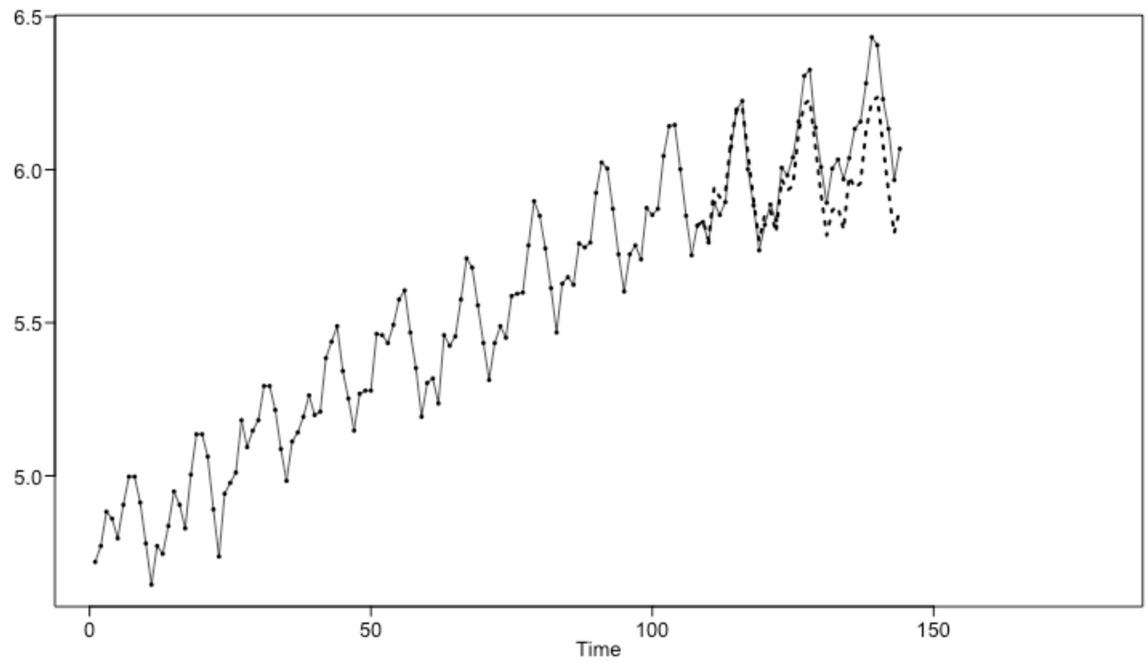
$$(1 - B)(1 - B^{12})(X_t - \mu) = (1 - .4B)\ (1 - .6B^{12})a_t$$



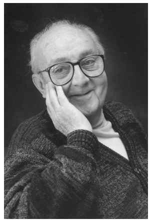
$$(1 + .36B + .05B^2 + .14B^3 + .11B^4 - .04B^5 - .09B^6 + .02B^7 - .02B^8 - .17B^9 - .03B^{10} + .10B^{11} + .38B^{12})(1 - B)(1 - B^{12})(X_t - \mu) = a_t$$

Airline Competition Among the Greats!

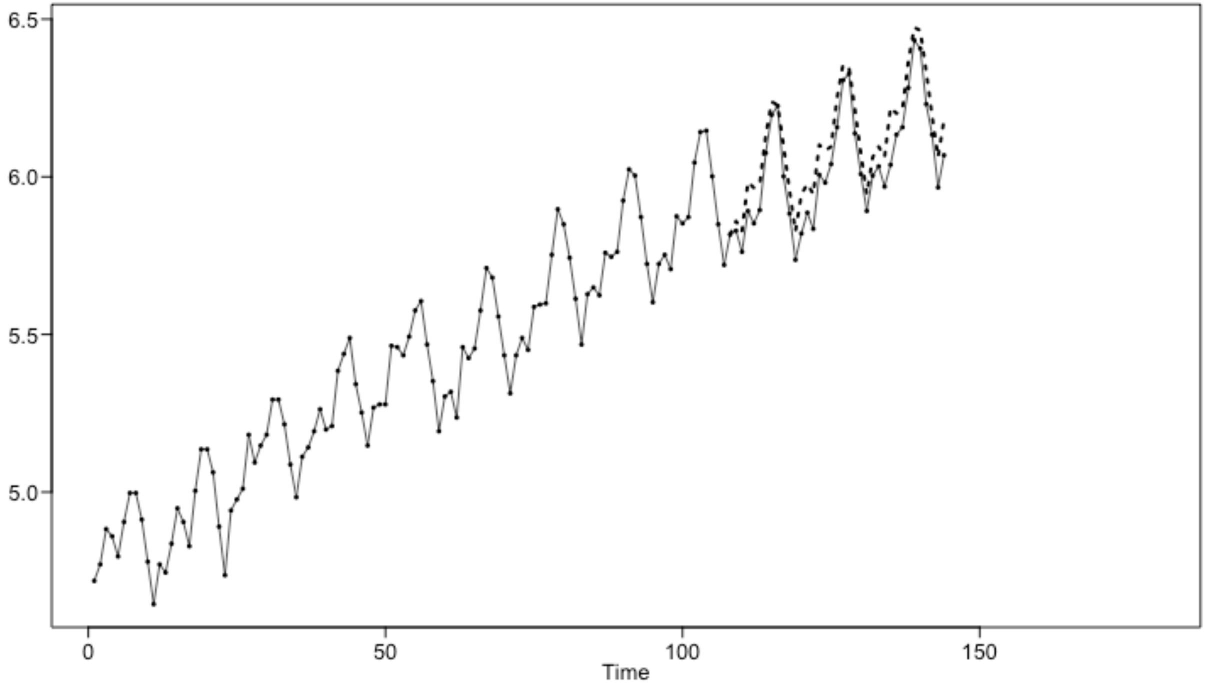
$$(1 - .74B)(1 + .38B^{12})(1 - B^{12})(X_t - \mu) = a_t$$



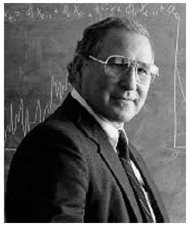
Airline Competition Among the Greats!



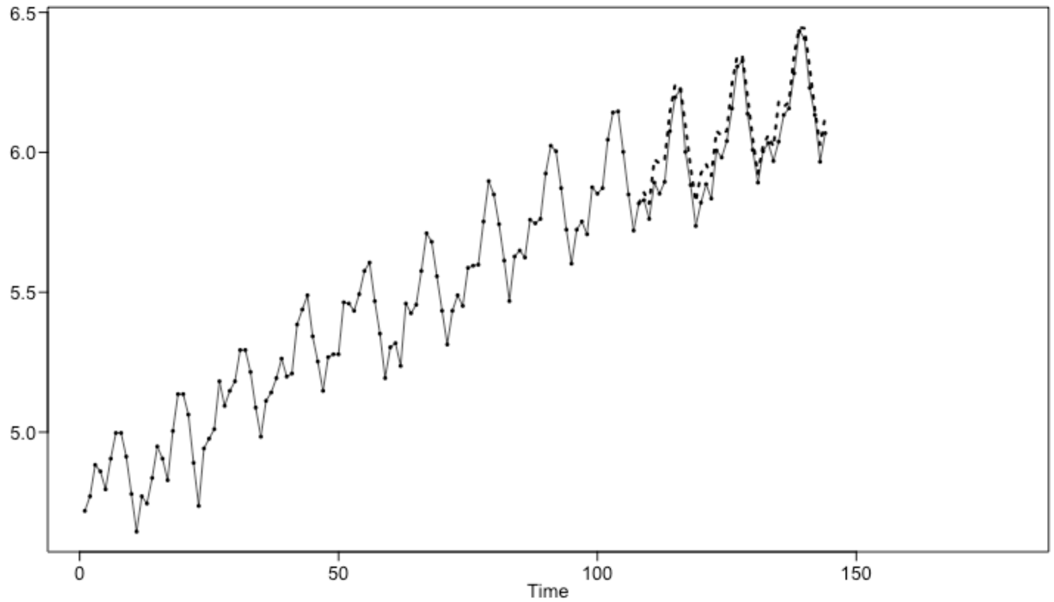
$$(1 - B)(1 - B^{12})(X_t - \mu) = (1 - .4B)\,(1 - .6B^{12})a_t$$



Airline Competition Among the Greats!



$$(1 + .36B + .05B^2 + .14B^3 + .11B^4 - .04B^5 - .09B^6 + .02B^7 - .02B^8 - .17B^9 - .03B^{10} + .10B^{11} + .38B^{12})(1 - B)(1 - B^{12})(X_t - \mu) = a_t$$

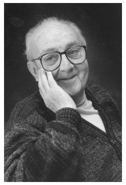


Airline Competition Among the Greats!

Now we will compare the actual models they published on the basis of the ASE calculated from forecasts of the last 36 years.

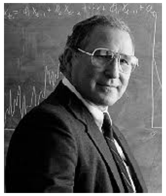
$$(1 - .74B)(1 + .38B^{12})(1 - B^{12})(X_t - \mu) = a_t$$

```
> Parzen = fore.aruma.wge(airlog, d = 0, s = 12, phi = c(.74,0,0,0,0,0,0,0,0,0,0,.38,-.2812),n.ahead = 36,lastn = TRUE, limits = FALSE)
> PARZEN_ASE = mean((airlog[(144-36+1):144] - Parzen$f)^2)
> PARZEN_ASE
[1] 0.01252636
```



$$(1 - B)(1 - B^{12})(X_t - \mu) = (1 - .4B)\ (1 - .6B^{12})a_t$$

```
> Box = fore.aruma.wge(airlog,d = 1, s = 12, theta = c(.4,0,0,0,0,0,0,0,0,0,0,.6,-.24),n.ahead = 36,lastn = TRUE, limits = FALSE)
> BOX_ASE = mean((airlog[(144-36+1):144] - Box$f)^2)
> BOX_ASE
[1] 0.006903242
```



$$(1 + .36B + .05B^2 + .14B^3 + .11B^4 - .04B^5 - .09B^6 + .02B^7 - .02B^8 - .17B^9 - .03B^{10} + .10B^{11} + .38B^{12})(1 - B)(1 - B^{12})(X_t - \mu) = a_t$$

?

DataScience@SMU

DataScience@SMU

Competition Conclusion

Airline Competition Among the Greats!

Now we will compare the actual models they published on the basis of the ASE calculated from forecasts of the last 36 years.

AIC

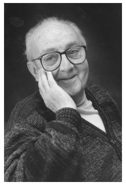
$$(1 - .74B)(1 + .38B^{12})(1 - B^{12})(X_t - \mu) = a_t$$

```
> Parzen$value
[1] -6.465559
```

```
> Parzen = fore.aruma.wge(air
> PARZEN_ASE = mean((airlog[(
> PARZEN_ASE
[1] 0.01252636
```

```
stn = TRUE, limits = FALSE)
```

Which model is the most useful?



$$(1 - B)(1$$

```
> Box$value
[1] -6.499275
```

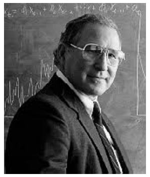
All of these professors would agree that it is nearly certain that none of these are the "right" model!

```
> Box = fore.aruma.wge(airl
> BOX_ASE = mean((airlog[(14
> BOX_ASE
[1] 0.006903242
```

```
tn = TRUE, limits = FALSE)
```

as well.

(BIC, AICc, RMSE, etc.)



$$(1 + \quad B^7 - .02B^8$$

$$- .17 \quad \mu) = a_t$$

```
> Woodward = fore.aruma.wge(airlog,d = 1, s = 12, phi = c(-.36,-.05,-.14,-.11,.04,.09,-.02, .02,.17,.03,-.10,-.38),n.ahead = 36,lastn = TRUE, limits = FALSE)
> WOODWARD_ASE = mean((airlog[(144-36+1):144] - Woodward$f)^2)
> WOODWARD_ASE
[1] 0.004185726
```

```
> Woodward$value
[1] -6.423649
```

DataScience@SMU

Forecasting Signal-Plus-Noise

Signal-Plus-Noise Forecasts

$$X_t = s_t + Z_t$$

tswge allows two cases:

$$s_t = b_0 + b_1 t$$

$$s_t = A_1 + A_2 \cos(2\;\; f_0 t + A_3) = A_1 + C_1 \cos(2\;\; f_0 t) + C_2 \sin(2\;\; f_0 t)$$

Forecasting Strategy (Linear Case)

Fit b_0 and b_1 using least squares (or use true values)

Calculate $\hat{Z}_t = X_t \quad b_0 \quad b_1 t$

Fit AR(p) to $\hat{Z}_t$ (covered in next unit)

Find forecasts $\hat{Z}_{t_0}(\ell)$

Forecasts for $X_{t_0+\ell}$ are $\hat{X}_{t_0}(\ell) = b_0 + b_1 t + \hat{Z}_{t_0}(\ell)$

Forecast limits for $\hat{X}_{t_0}(\ell)$ are

$$b_0 + b_1 t + \hat{Z}_{t_0}(\ell) \pm 1.96\ \hat{\ }_a \left(\ {}^{\ell}_{k=0} \quad {}^{2}_{k} \right)^2$$

where $\hat{\ }_{\mathrm{a}}$ and $\ _k$ are based on AR(p) model fit to $\hat{Z}_t$

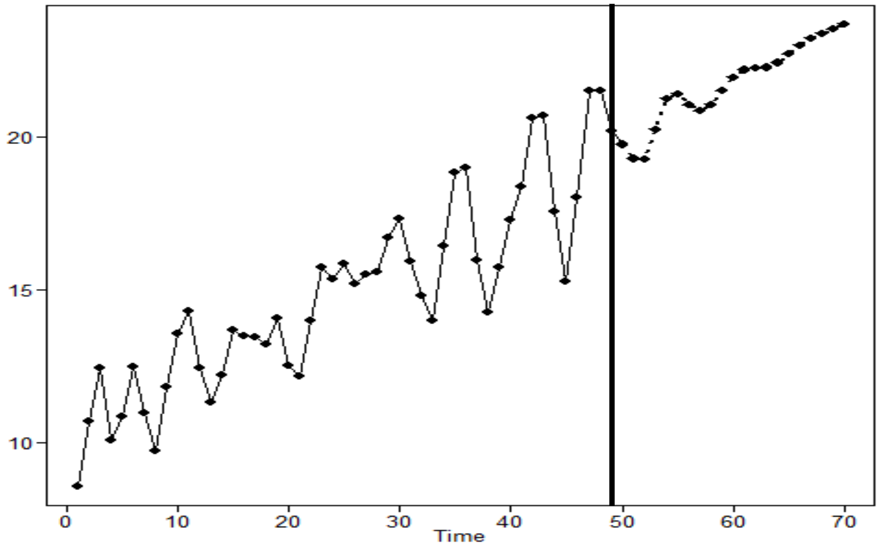
In tswge:

fore.sigplusnoise.wge(x, linear = TRUE, freq = 0, max.p = 5, n.ahead = 10, lastn = FALSE, plot = TRUE, limits = TRUE)

x=gen.sigplusnoise.wge(n=50,b0=10,b1=.2, phi=c(.8,-.6))

#

xfore=fore.sigplusnoise.wge(x,linear=TRUE,n.ahead=20,lastn=F,limits=F)



Notice that the realization has a line with cyclic noise. The forecasts for early lags forecast the cyclic behavior, while for longer steps ahead, they trend to the underlying line.

Closing Remarks

Final Note

We have learned how to forecast from a **given** ARMA, ARIMA, or signal-plus-noise model.

Important point

- The model selected makes a big difference in the forecasts.
- We saw this in the airline data: We needed a factor of $(1-B)$ along with $(1-B^{12})$ to forecast the trend to continue.
- **How do we fit the parameters to these models?**
- **How do we decide what model to use?**
    - *Upcoming Units!*

DataScience@SMU