Forecasting: principles and practice

Rob J Hyndman

2.5 Seasonal ARIMA models

Outline

Backshift notation

A very useful notational device is the backward shift operator, B, which is used as follows:

$$By_t = y_{t-1} \ .$$

In other words, B, operating on y_t, has the effect of **shifting the data back one period**. Two applications of B to y_t **shifts the data back two periods:**

$$B(By_t) = B^2 y_t = y_{t-2} \ .$$

For monthly data, if we wish to shift attention to "the same month last year," then B^{12} is used, and the notation is $B^{12} y_t = y_{t-12}$.

Backshift notation

A very useful notational device is the backward shift operator, B, which is used as follows:

$$By_t = y_{t-1}\ .$$

In other words, B, operating on y_t, has the effect of **shifting the data back one period**. Two applications of B to y_t **shifts the data back two periods**:

$$B(By_t) = B^2 y_t = y_{t-2}\ .$$

For monthly data, if we wish to shift attention to "the same month last year," then B^{12} is used, and the notation is $B^{12}y_t = y_{t-12}$.

Backshift notation

A very useful notational device is the backward shift operator, B, which is used as follows:

$$By_t = y_{t-1}\ .$$

In other words, B, operating on y_t, has the effect of **shifting the data back one period**. Two applications of B to y_t **shifts the data back two periods**:

$$B(By_t) = B^2y_t = y_{t-2}\ .$$

For monthly data, if we wish to shift attention to "the same month last year," then B^{12} is used, and the notation is $B^{12}y_t = y_{t-12}$.

Backshift notation

A very useful notational device is the backward shift operator, B, which is used as follows:

$$By_t = y_{t-1} \; .$$

In other words, B, operating on y_t, has the effect of **shifting the data back one period**. Two applications of B to y_t **shifts the data back two periods**:

$$B(By_t) = B^2 y_t = y_{t-2} \; .$$

For monthly data, if we wish to shift attention to "the same month last year," then B^{12} is used, and the notation is $B^{12} y_t = y_{t-12}$.

Backshift notation

- First difference: $1 - B$.
- Double difference: $(1 - B)^2$.
- dth-order difference: $(1 - B)^d y_t$.
- Seasonal difference: $1 - B^m$.
- Seasonal difference followed by a first difference: $(1 - B)(1 - B^m)$.
- Multiply terms together together to see the combined effect:

$$\begin{aligned}(1 - B)(1 - B^m)y_t &= (1 - B - B^m + B^{m+1})y_t \\ &= y_t - y_{t-1} - y_{t-m} + y_{t-m-1}.\end{aligned}$$

Backshift notation for ARIMA

- ARMA model:

$$\begin{aligned} y_t &= c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + e_t + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} \\ &= c + \phi_1 B y_t + \cdots + \phi_p B^p y_t + e_t + \theta_1 B e_t + \cdots + \theta_q B^q e_t \\ \phi(B) y_t &= c + \theta(B) e_t \end{aligned}$$

where $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$

and $\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$.

- ARIMA(1,1,1) model:

$$\underset{\substack{\uparrow \\ \text{AR(1)}}}{(1 - \phi_1 B)} \quad \underset{\substack{\uparrow \\ \text{First difference}}}{(1 - B) y_t} \quad = \quad c + \underset{\substack{\uparrow \\ \text{MA(1)}}}{(1 + \theta_1 B) e_t}$$

Backshift notation for ARIMA

- ARMA model:

$$\begin{aligned} y_t &= c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + e_t + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} \\ &= c + \phi_1 B y_t + \cdots + \phi_p B^p y_t + e_t + \theta_1 B e_t + \cdots + \theta_q B^q e_t \\ \phi(B) y_t &= c + \theta(B) e_t \end{aligned}$$

where $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$

and $\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$.

- ARIMA(1,1,1) model:

$$\underbrace{(1 - \phi_1 B)}_{\uparrow \text{ AR(1)}} \quad \underbrace{(1 - B) y_t}_{\uparrow \text{ First difference}} \quad = \quad c + \underbrace{(1 + \theta_1 B) e_t}_{\uparrow \text{ MA(1)}}$$

Backshift notation for ARIMA

- ARIMA(p, d, q) model:

$$\underbrace{(1 - \phi_1 B - \cdots - \phi_p B^p)}_{\uparrow \atop \text{AR}(p)} \; \underbrace{(1 - B)^d}_{\uparrow \atop d \text{ differences}} y_t = c + \underbrace{(1 + \theta_1 B + \cdots + \theta_q B^q)}_{\uparrow \atop \text{MA}(q)} e_t$$

Outline

Seasonal ARIMA models

ARIMA	$\underbrace{(p, d, q)}$	$\underbrace{(P, D, Q)_m}$
	$\uparrow$ Non-seasonal part of the model	$\uparrow$ Seasonal part of of the model

where m = number of observations per year.

Seasonal ARIMA models

E.g., ARIMA$(1, 1, 1)(1, 1, 1)_4$ model (without constant)

$$\underbrace{(1-\phi_1 B)}_{\text{Non-seasonal AR(1)}}\underbrace{(1-\Phi_1 B^4)}_{\text{Seasonal AR(1)}}\underbrace{(1-B)}_{\text{Non-seasonal difference}}\underbrace{(1-B^4)}_{\text{Seasonal difference}}y_t = \underbrace{(1+\theta_1 B)}_{\text{Non-seasonal MA(1)}}\underbrace{(1+\Theta_1 B^4)}_{\text{Seasonal MA(1)}}e_t.$$

Seasonal ARIMA models

E.g., ARIMA$(1, 1, 1)(1, 1, 1)_4$ model (without constant)

$$\underbrace{(1-\phi_1 B)}_{\text{Non-seasonal AR(1)}}\underbrace{(1-\Phi_1 B^4)}_{\text{Seasonal AR(1)}}\underbrace{(1-B)}_{\text{Non-seasonal difference}}\underbrace{(1-B^4)}_{\text{Seasonal difference}}y_t = \underbrace{(1+\theta_1 B)}_{\text{Non-seasonal MA(1)}}\underbrace{(1+\Theta_1 B^4)}_{\text{Seasonal MA(1)}}e_t.$$

Seasonal ARIMA models

E.g., ARIMA$(1, 1, 1)(1, 1, 1)_4$ model (without constant)

$$\underbrace{(1-\phi_1 B)}_{\begin{pmatrix}\text{Non-seasonal}\\ \text{AR(1)}\end{pmatrix}}\underbrace{(1-\Phi_1 B^4)}_{\begin{pmatrix}\text{Seasonal}\\ \text{AR(1)}\end{pmatrix}}\underbrace{(1-B)}_{\begin{pmatrix}\text{Non-seasonal}\\ \text{difference}\end{pmatrix}}\underbrace{(1-B^4)}_{\begin{pmatrix}\text{Seasonal}\\ \text{difference}\end{pmatrix}}y_t = \underbrace{(1+\theta_1 B)}_{\begin{pmatrix}\text{Non-seasonal}\\ \text{MA(1)}\end{pmatrix}}\underbrace{(1+\Theta_1 B^4)}_{\begin{pmatrix}\text{Seasonal}\\ \text{MA(1)}\end{pmatrix}}e_t.$$

Seasonal ARIMA models

E.g., ARIMA$(1, 1, 1)(1, 1, 1)_4$ model (without constant)

$$(1-\phi_1 B)(1-\Phi_1 B^4)(1-B)(1-B^4)y_t = (1+\theta_1 B)(1+\Theta_1 B^4)e_t.$$

All the factors can be multiplied out and the general model written as follows:

$$\begin{aligned}
y_t &= (1+\phi_1)y_{t-1} - \phi_1 y_{t-2} + (1+\Phi_1)y_{t-4} \\
&\quad - (1+\phi_1+\Phi_1+\phi_1\Phi_1)y_{t-5} + (\phi_1+\phi_1\Phi_1)y_{t-6} \\
&\quad - \Phi_1 y_{t-8} + (\Phi_1+\phi_1\Phi_1)y_{t-9} - \phi_1\Phi_1 y_{t-10} \\
&\quad + e_t + \theta_1 e_{t-1} + \Theta_1 e_{t-4} + \theta_1\Theta_1 e_{t-5}.
\end{aligned}$$

Common ARIMA models

In the US Census Bureau uses the following models most often:

ARIMA$(0,1,1)(0,1,1)_m$	with log transformation
ARIMA$(0,1,2)(0,1,1)_m$	with log transformation
ARIMA$(2,1,0)(0,1,1)_m$	with log transformation
ARIMA$(0,2,2)(0,1,1)_m$	with log transformation
ARIMA$(2,1,2)(0,1,1)_m$	with no transformation

Seasonal ARIMA models

The seasonal part of an AR or MA model will be seen in the seasonal lags of the PACF and ACF.

ARIMA(0,0,0)(0,0,1)$_{12}$ will show:

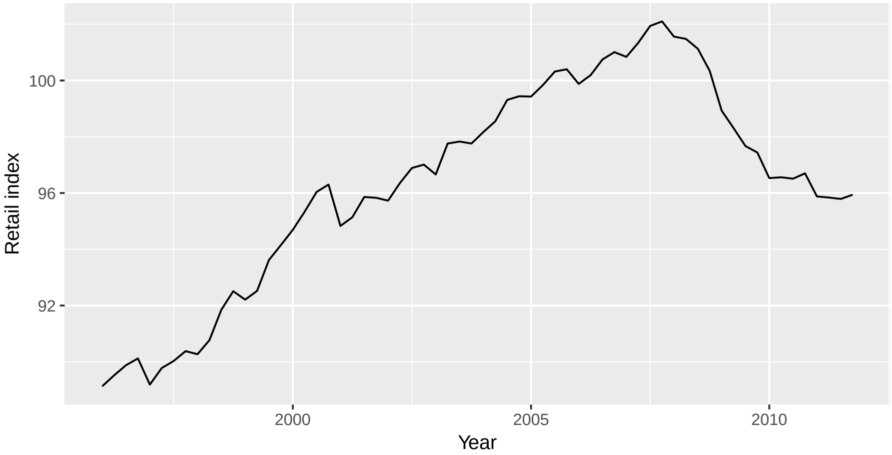
- a spike at lag 12 in the ACF but no other significant spikes.
- The PACF will show exponential decay in the seasonal lags; that is, at lags 12, 24, 36, ...

ARIMA(0,0,0)(1,0,0)$_{12}$ will show:

- exponential decay in the seasonal lags of the ACF
- a single significant spike at lag 12 in the PACF.

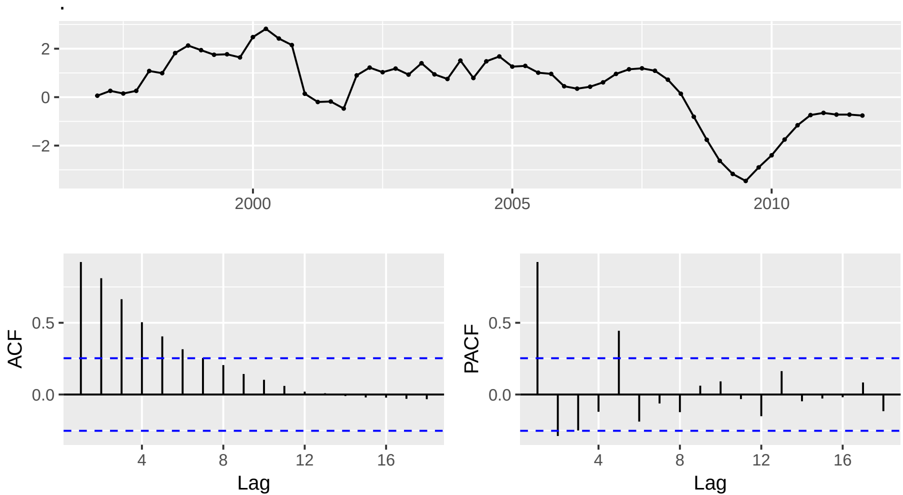
European quarterly retail trade

```r
autoplot(euretail) +
  xlab("Year") + ylab("Retail index")
```



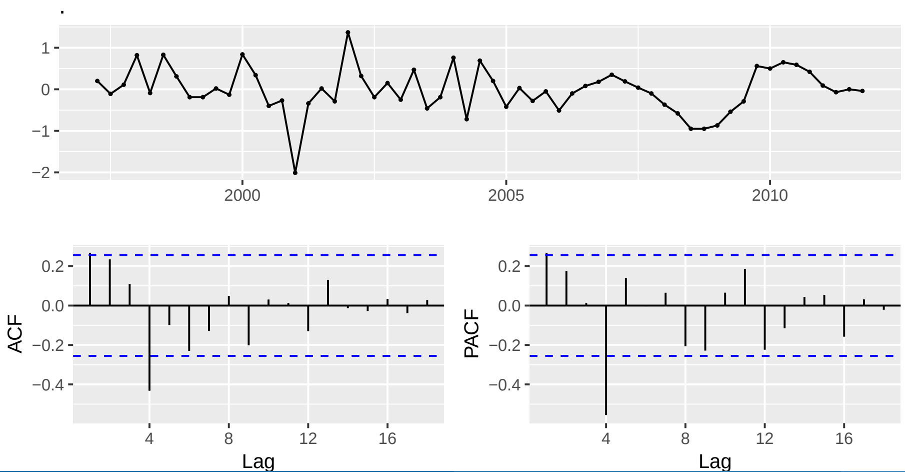
European quarterly retail trade

```r
euretail %>% diff(lag=4) %>% ggtsdisplay()
```



European quarterly retail trade

```r
euretail %>% diff(lag=4) %>% diff() %>%
  ggtsdisplay()
```

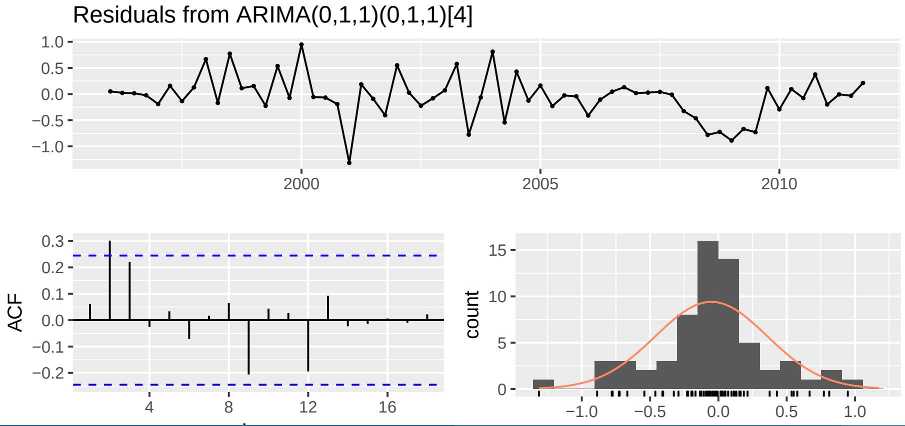


European quarterly retail trade

- $d = 1$ and $D = 1$ seems necessary.
- Significant spike at lag 1 in ACF suggests non-seasonal MA(1) component.
- Significant spike at lag 4 in ACF suggests seasonal MA(1) component.
- Initial candidate model: ARIMA$(0,1,1)(0,1,1)_4$.
- We could also have started with ARIMA$(1,1,0)(1,1,0)_4$.

European quarterly retail trade

```r
fit <- Arima(euretail, order=c(0,1,1),
  seasonal=c(0,1,1))
checkresiduals(fit)
```



European quarterly retail trade

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(0,1,1)(0,1,1)[4]
## Q* = 10.654, df = 6, p-value = 0.09968
##
## Model df: 2.   Total lags used: 8
```

European quarterly retail trade

- ACF and PACF of residuals show significant spikes at lag 2, and maybe lag 3.
- AICc of ARIMA$(0,1,2)(0,1,1)_4$ model is 74.36.
- AICc of ARIMA$(0,1,3)(0,1,1)_4$ model is 68.53.

```
fit <- Arima(euretail, order=c(0,1,3),
  seasonal=c(0,1,1))
checkresiduals(fit)
```

European quarterly retail trade

- ACF and PACF of residuals show significant spikes at lag 2, and maybe lag 3.
- AICc of ARIMA(0,1,2)(0,1,1)$_4$ model is 74.36.
- AICc of ARIMA(0,1,3)(0,1,1)$_4$ model is 68.53.

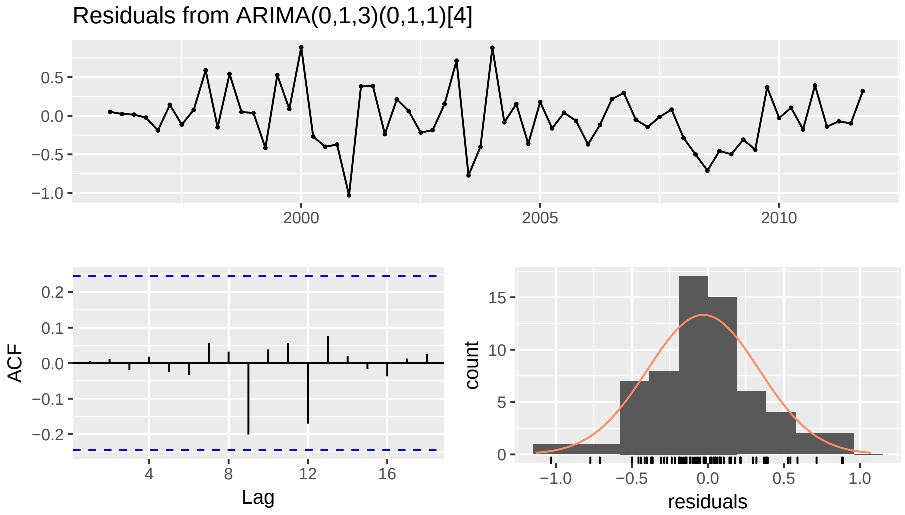
```
fit <- Arima(euretail, order=c(0,1,3),
  seasonal=c(0,1,1))
checkresiduals(fit)
```

European quarterly retail trade

```
## Series: euretail
## ARIMA(0,1,3)(0,1,1)[4]
##
## Coefficients:
##          ma1     ma2     ma3     sma1
##       0.2630  0.3694  0.4200  -0.6636
## s.e.  0.1237  0.1255  0.1294   0.1545
##
## sigma^2 estimated as 0.156:  log likelihood=-28.63
## AIC=67.26   AICc=68.39   BIC=77.65
```

European quarterly retail trade

```r
checkresiduals(fit)
```

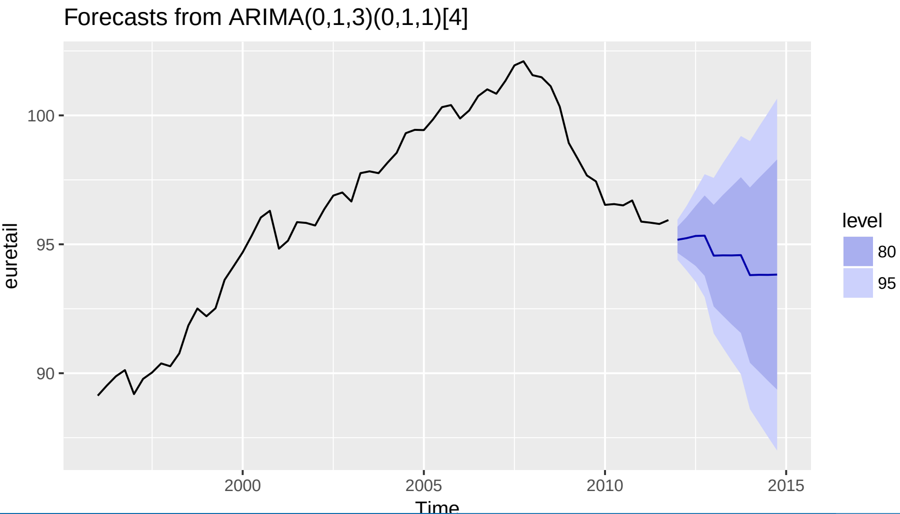


European quarterly retail trade

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(0,1,3)(0,1,1)[4]
## Q* = 0.51128, df = 4, p-value = 0.9724
##
## Model df: 4.   Total lags used: 8
```

European quarterly retail trade

```
autoplot(forecast(fit, h=12))
```



European quarterly retail trade

```r
auto.arima(euretail)
```

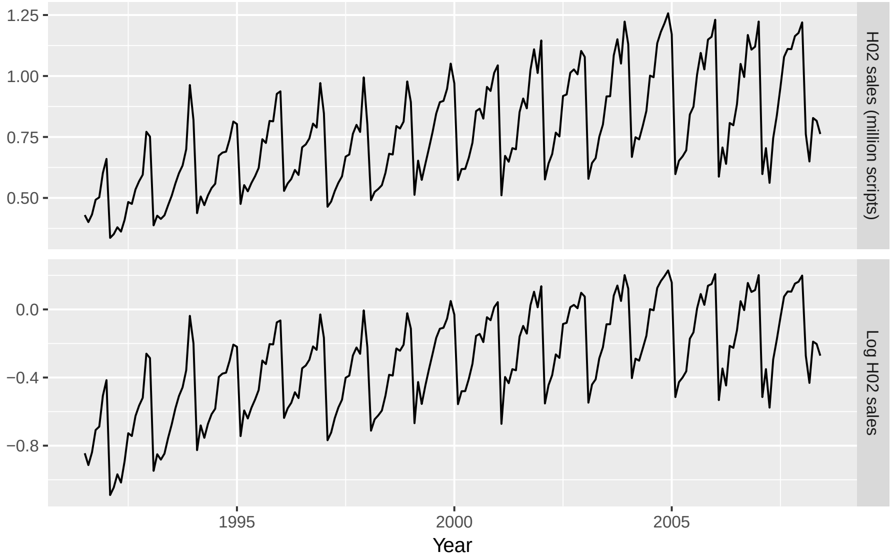
```
## Series: euretail
## ARIMA(1,1,2)(0,1,1)[4]
##
## Coefficients:
##          ar1      ma1     ma2     sma1
##       0.7362  -0.4663  0.2163  -0.8433
## s.e.  0.2243   0.1990  0.2101   0.1876
##
## sigma^2 estimated as 0.1587:  log likelihood=-29.62
## AIC=69.24   AICc=70.38   BIC=79.63
```

European quarterly retail trade

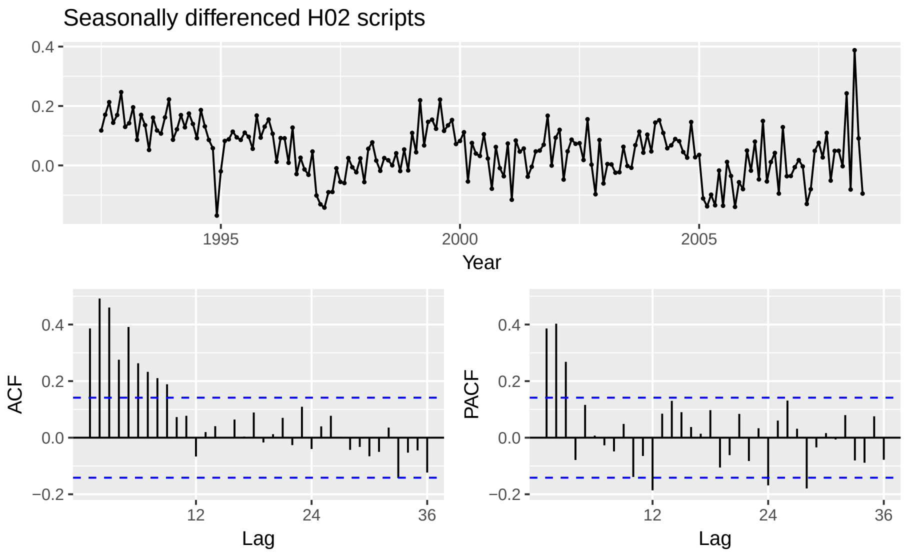
```r
auto.arima(euretail, stepwise=FALSE, approximation=FALSE)
```

```
## Series: euretail
## ARIMA(0,1,3)(0,1,1)[4]
##
## Coefficients:
##          ma1     ma2     ma3     sma1
##       0.2630  0.3694  0.4200  -0.6636
## s.e.  0.1237  0.1255  0.1294   0.1545
##
## sigma^2 estimated as 0.156:  log likelihood=-28.63
## AIC=67.26   AICc=68.39   BIC=77.65
```

Cortecosteroid drug sales



Cortecosteroid drug sales



Cortecosteroid drug sales

- Choose $D = 1$ and $d = 0$.
- Spikes in PACF at lags 12 and 24 suggest seasonal AR(2) term.
- Spikes in PACF suggests possible non-seasonal AR(3) term.
- Initial candidate model: ARIMA$(3,0,0)(2,1,0)_{12}$.

Cortecosteroid drug sales

Model	AICc
ARIMA$(3,0,0)(2,1,0)_{12}$	-475.12
ARIMA$(3,0,1)(2,1,0)_{12}$	-476.31
ARIMA$(3,0,2)(2,1,0)_{12}$	-474.88
ARIMA$(3,0,1)(1,1,0)_{12}$	-463.40
ARIMA$(3,0,1)(0,1,1)_{12}$	-483.67
ARIMA$(3,0,1)(0,1,2)_{12}$	-485.48
ARIMA$(3,0,1)(1,1,1)_{12}$	-484.25

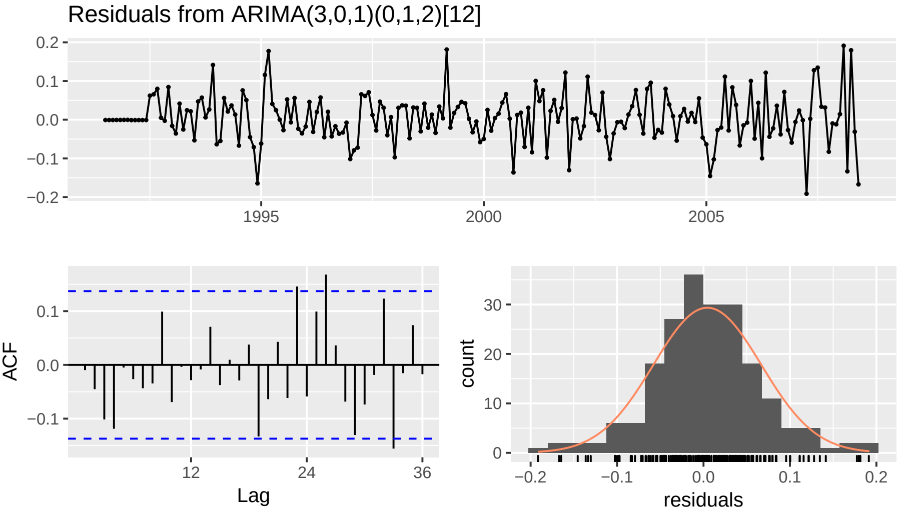
Cortecosteroid drug sales

```r
(fit <- Arima(h02, order=c(3,0,1), seasonal=c(0,1,2),
  lambda=0))
```

```
## Series: h02
## ARIMA(3,0,1)(0,1,2)[12]
## Box Cox transformation: lambda= 0
##
## Coefficients:
##           ar1     ar2     ar3     ma1     sma1     sma2
##       -0.1603  0.5481  0.5678  0.3827  -0.5222  -0.1768
## s.e.   0.1636  0.0878  0.0942  0.1895   0.0861   0.0872
##
## sigma^2 estimated as 0.004278:  log likelihood=250.04
## AIC=-486.08   AICc=-485.48   BIC=-463.28
```

Cortecosteroid drug sales

```r
checkresiduals(fit)
```



Cortecosteroid drug sales

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(3,0,1)(0,1,2)[12]
## Q* = 23.663, df = 18, p-value = 0.1664
##
## Model df: 6.   Total lags used: 24
```

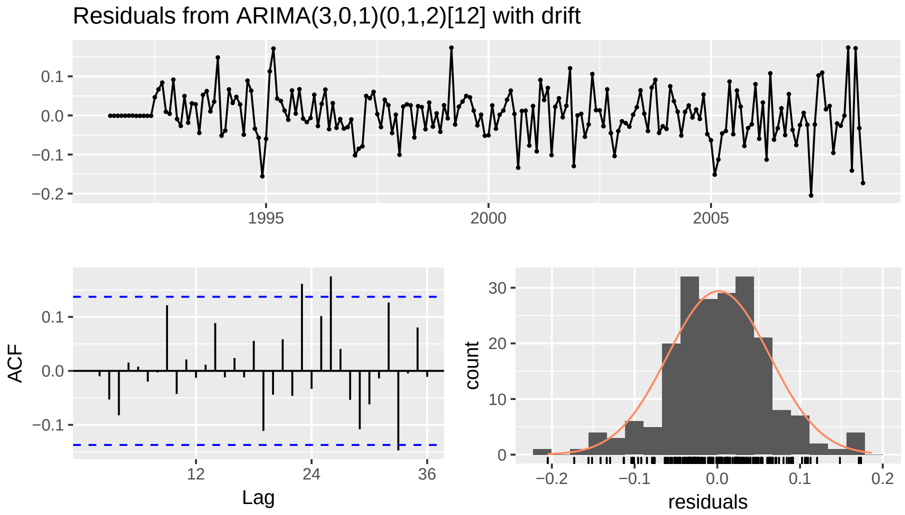
Cortecosteroid drug sales

```
(fit <- auto.arima(h02, lambda=0, d=0, D=1, max.order=9,
  stepwise=FALSE, approximation=FALSE))
```

```
## Series: h02
## ARIMA(3,0,1)(0,1,2)[12] with drift
## Box Cox transformation: lambda= 0
##
## Coefficients:
##           ar1     ar2     ar3     ma1     sma1     sma2   drift
##       -0.2653  0.5011  0.5394  0.4572  -0.5031  -0.2030  0.0038
## s.e.   0.1691  0.0813  0.0848  0.1904   0.0847   0.0871  0.0009
##
## sigma^2 estimated as 0.004176:  log likelihood=252.99
## AIC=-489.99   AICc=-489.2   BIC=-463.93
```

Cortecosteroid drug sales

```r
checkresiduals(fit)
```



Cortecosteroid drug sales

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(3,0,1)(0,1,2)[12] with
## Q* = 19.369, df = 17, p-value = 0.3078
##
## Model df: 7.   Total lags used: 24
```

Cortecosteroid drug sales

Training data: July 1991 to June 2006

Test data: July 2006–June 2008

```
getrmse <- function(x,h,...)
{
  train.end <- time(x)[length(x)-h]
  test.start <- time(x)[length(x)-h+1]
  train <- window(x,end=train.end)
  test <- window(x,start=test.start)
  fit <- Arima(train,...)
  fc <- forecast(fit,h=h)
  return(accuracy(fc,test)[2,"RMSE"])
}
getrmse(h02,h=24,order=c(3,0,0),seasonal=c(2,1,0),lambda=0)
getrmse(h02,h=24,order=c(3,0,1),seasonal=c(2,1,0),lambda=0)
getrmse(h02,h=24,order=c(3,0,2),seasonal=c(2,1,0),lambda=0)
getrmse(h02,h=24,order=c(3,0,1),seasonal=c(1,1,0),lambda=0)
getrmse(h02,h=24,order=c(3,0,1),seasonal=c(0,1,1),lambda=0)
getrmse(h02,h=24,order=c(3,0,1),seasonal=c(0,1,2),lambda=0)
```

Cortecosteroid drug sales

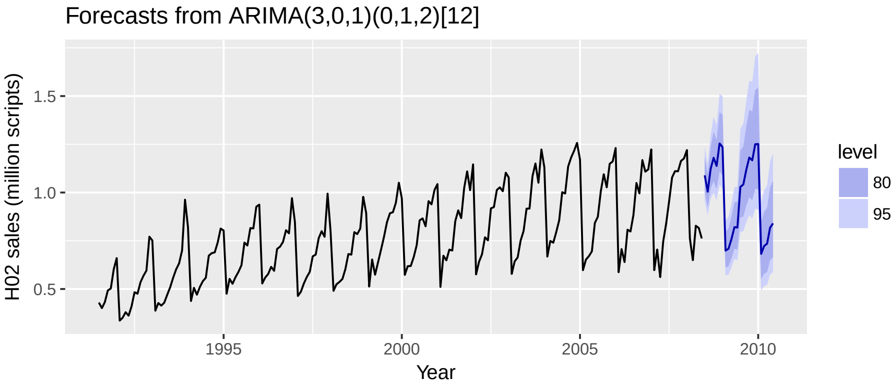
Model	RMSE
ARIMA(3,0,0)(2,1,0)[12]	0.0661
ARIMA(3,0,1)(2,1,0)[12]	0.0646
ARIMA(3,0,2)(2,1,0)[12]	0.0645
ARIMA(3,0,1)(1,1,0)[12]	0.0679
ARIMA(3,0,1)(0,1,1)[12]	0.0644
ARIMA(3,0,1)(0,1,2)[12]	0.0622
ARIMA(3,0,1)(1,1,1)[12]	0.0630
ARIMA(4,0,3)(0,1,1)[12]	0.0648
ARIMA(3,0,3)(0,1,1)[12]	0.0639
ARIMA(4,0,2)(0,1,1)[12]	0.0648
ARIMA(3,0,2)(0,1,1)[12]	0.0644
ARIMA(2,1,3)(0,1,1)[12]	0.0634
ARIMA(2,1,4)(0,1,1)[12]	0.0632
ARIMA(2,1,5)(0,1,1)[12]	0.0640

Cortecosteroid drug sales

- Models with lowest AICc values tend to give slightly better results than the other models.
- AICc comparisons must have the same orders of differencing. But RMSE test set comparisons can involve any models.
- No model passes all the residual tests.
- Use the best model available, even if it does not pass all tests.
- In this case, the ARIMA(3,0,1)(0,1,2)$_{12}$ has the lowest RMSE value and the best AICc value for models with fewer than 6 parameters.

Cortecosteroid drug sales

```
fit <- Arima(h02, order=c(3,0,1), seasonal=c(0,1,2),
  lambda=0)
autoplot(forecast(fit)) +
  ylab("H02 sales (million scripts)") + xlab("Year")
```



Outline

ARIMA vs ETS

- Myth that ARIMA models are more general than exponential smoothing.
- Linear exponential smoothing models all special cases of ARIMA models.
- Non-linear exponential smoothing models have no equivalent ARIMA counterparts.
- Many ARIMA models have no exponential smoothing counterparts.
- ETS models all non-stationary. Models with seasonality or non-damped trend (or both) have two unit roots; all other models have one unit root.

Equivalences

ETS model	ARIMA model	Parameters
ETS(A,N,N)	ARIMA(0,1,1)	$\theta_1 = \alpha - 1$
ETS(A,A,N)	ARIMA(0,2,2)	$\theta_1 = \alpha + \beta - 2$
		$\theta_2 = 1 - \alpha$
ETS(A,A,N)	ARIMA(1,1,2)	$\phi_1 = \phi$
		$\theta_1 = \alpha + \phi\beta - 1 - \phi$
		$\theta_2 = (1 - \alpha)\phi$
ETS(A,N,A)	ARIMA(0,0,m)(0,1,0)$_m$	
ETS(A,A,A)	ARIMA(0,1,$m+1$)(0,1,0)$_m$	
ETS(A,A,A)	ARIMA(1,0,$m+1$)(0,1,0)$_m$	

Outline

Lab Session 12