2.5 Seasonal ARIMA models
1. Backshift notation reviewed
2. Seasonal ARIMA models
3. ARIMA vs ETS
4. Lab session 12
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 $B^{12} y_t = y_{t-12}$. 

Forecasting: principles and practice
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 $B^{12}y_t = y_{t-12}$. 
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 $B^{12}y_t = y_{t-12}$. 
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

- First difference: $1 - B$.
- Double difference: $(1 - B)^2$.
- $d$th-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 to see the combined effect:

$$
(1 - B)(1 - B^m)y_t = (1 - B - B^m + B^{m+1})y_t
$$

$$
= y_t - y_{t-1} - y_{t-m} + y_{t-m-1}.
$$
Backshift notation for ARIMA

**ARMA model:**

\[ y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + e_t + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} \]

\[ = c + \phi_1 B y_t + \cdots + \phi_p B^p y_t + e_t + \theta_1 B e_t + \cdots + \theta_q B^q e_t \]

\[ \phi(B) y_t = c + \theta(B) e_t \]

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:**

\[ (1 - \phi_1 B) (1 - B) y_t = c + (1 + \theta_1 B) e_t \]

\[ \uparrow \quad \uparrow \quad \uparrow \]

AR(1) First difference MA(1)
Backshift notation for ARIMA

ARMA model:

\[ y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + e_t + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} \]

\[ = c + \phi_1 B y_t + \cdots + \phi_p B^p y_t + e_t + \theta_1 B e_t + \cdots + \theta_q B^q e_t \]

\[ \phi(B)y_t = c + \theta(B)e_t \]

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:

\[ (1 - \phi_1 B) (1 - B)y_t = c + (1 + \theta_1 B)e_t \]

↑

AR(1) First difference

↑

↑

MA(1)
Backshift notation for ARIMA

- ARIMA($p$, $d$, $q$) model:

\[
(1 - \phi_1 B - \cdots - \phi_p B^p) (1 - B)^d y_t = c + (1 + \theta_1 B + \cdots + \theta_q B^q) e_t
\]

\[
\uparrow \quad \uparrow \quad \uparrow
\]

AR($p$) \hspace{2cm} d \text{ differences} \hspace{2cm} MA(q)

Forecasting: principles and practice
1. Backshift notation reviewed
2. Seasonal ARIMA models
3. ARIMA vs ETS
4. Lab session 12
Seasonal ARIMA models

ARIMA \( (p, d, q) \) \( \uparrow \) \( (P, D, Q)_m \)

Non-seasonal part of the model \( \uparrow \) Seasonal part 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)

\[(1 - \phi_1 B)(1 - \Phi_1 B^4)(1 - B)(1 - B^4)y_t = (1 + \theta_1 B)(1 + \Theta_1 B^4)e_t.\]
Seasonal ARIMA models

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

\[
(1 - \phi_1 B)(1 - \Phi_1 B^4)(1 - B)(1 - B^4)y_t = (1 + \theta_1 B)(1 + \Theta_1 B^4)e_t.
\]
Seasonal ARIMA models

E.g., ARIMA(1, 1, 1)(1, 1, 1) \_4 \text{ model (without constant)}

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

\[y_t = (1 + \phi_1)y_{t-1} - \phi_1 y_{t-2} + (1 + \Phi_1)y_{t-4} - \Phi_1 y_{t-8} + (\Phi_1 + \phi_1 \Phi_1)y_{t-9} - \phi_1 \Phi_1 y_{t-10} + e_t + \theta_1 e_{t-1} + \Theta_1 e_{t-4} + \theta_1 \Theta_1 e_{t-5}.\]
Common ARIMA models

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

<table>
<thead>
<tr>
<th>Model Description</th>
<th>Transformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA(0,1,1)(0,1,1)_m</td>
<td>with log transformation</td>
</tr>
<tr>
<td>ARIMA(0,1,2)(0,1,1)_m</td>
<td>with log transformation</td>
</tr>
<tr>
<td>ARIMA(2,1,0)(0,1,1)_m</td>
<td>with log transformation</td>
</tr>
<tr>
<td>ARIMA(0,2,2)(0,1,1)_m</td>
<td>with log transformation</td>
</tr>
<tr>
<td>ARIMA(2,1,2)(0,1,1)_m</td>
<td>with no transformation</td>
</tr>
</tbody>
</table>
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

euretail %>% diff(lag=4) %>% ggtsdisplay()
euretail %>% diff(lag=4) %>% diff() %>% ggtsdisplay()
$d = 1$ 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($1, 1, 0)(1, 1, 0)_4$. 
fit <- Arima(euretail, order=c(0,1,1), seasonal=c(0,1,1))
checkresiduals(fit)

Residuals from ARIMA(0,1,1)(0,1,1)[4]

ACF

Lag

0 5 10 15

residuals
count

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
## 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
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.

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

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

Residuals from ARIMA(0,1,3)(0,1,1)[4]

ACF

Lag

count

residuals
## 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

\texttt{autoplot(\texttt{forecast(fit, h=12)})}

Forecasts from ARIMA\((0,1,3)(0,1,1)[4]\)
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
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

Forecasting: principles and practice
Seasonal ARIMA models
Corticosteroid drug sales

Seasonally differenced H02 scripts

Forecasting: principles and practice
Seasonal ARIMA models
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: $\text{ARIMA}(3,0,0)(2,1,0)_{12}$.
Cortecosteroid drug sales

<table>
<thead>
<tr>
<th>Model</th>
<th>AICc</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA(3,0,0)(2,1,0)$_{12}$</td>
<td>-475.12</td>
</tr>
<tr>
<td>ARIMA(3,0,1)(2,1,0)$_{12}$</td>
<td>-476.31</td>
</tr>
<tr>
<td>ARIMA(3,0,2)(2,1,0)$_{12}$</td>
<td>-474.88</td>
</tr>
<tr>
<td>ARIMA(3,0,1)(1,1,0)$_{12}$</td>
<td>-463.40</td>
</tr>
<tr>
<td>ARIMA(3,0,1)(0,1,1)$_{12}$</td>
<td>-483.67</td>
</tr>
<tr>
<td>ARIMA(3,0,1)(0,1,2)$_{12}$</td>
<td>-485.48</td>
</tr>
<tr>
<td>ARIMA(3,0,1)(1,1,1)$_{12}$</td>
<td>-484.25</td>
</tr>
</tbody>
</table>
library(forecast)

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

checkresiduals(fit)

Residuals from ARIMA(3,0,1)(0,1,2)[12]

<table>
<thead>
<tr>
<th>Lag</th>
<th>12</th>
<th>24</th>
<th>36</th>
</tr>
</thead>
<tbody>
<tr>
<td>count</td>
<td>30</td>
<td>20</td>
<td>10</td>
</tr>
</tbody>
</table>

ACF

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
## 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

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

```r
## 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
## sigma^2 estimated as 0.004176: log likelihood=252.99
## AIC=-489.99 AICc=-489.2 BIC=-463.93
```
Corticosteroid drug sales

checkresiduals(fit)

Residuals from ARIMA(3,0,1)(0,1,2)[12] with drift

Ljung-Box test

\[
\text{data: Residuals from ARIMA(3,0,1)(0,1,2)[12] with drift}
\]

\[
Q^* = 19.369, \quad \text{df} = 17, \quad p-value = 0.3078
\]

Model df: 7. Total lags used: 24
## Ljung-Box test

## data: Residuals from ARIMA(3,0,1)(0,1,2)[12] with drift

\[ Q^* = 19.369, \text{ df } = 17, \text{ p-value } = 0.3078 \]

## Model df: 7. Total lags used: 24
Corticosteroid drug sales

Training data: July 1991 to June 2006
Test data: July 2006–June 2008

```r
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)
```

Forecasting: principles and practice
Seasonal ARIMA models
Corticosteroid drug sales

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA(3,0,0)(2,1,0)[12]</td>
<td>0.0661</td>
</tr>
<tr>
<td>ARIMA(3,0,1)(2,1,0)[12]</td>
<td>0.0646</td>
</tr>
<tr>
<td>ARIMA(3,0,2)(2,1,0)[12]</td>
<td>0.0645</td>
</tr>
<tr>
<td>ARIMA(3,0,1)(1,1,0)[12]</td>
<td>0.0679</td>
</tr>
<tr>
<td>ARIMA(3,0,1)(0,1,1)[12]</td>
<td>0.0644</td>
</tr>
<tr>
<td>ARIMA(3,0,1)(0,1,2)[12]</td>
<td>0.0622</td>
</tr>
<tr>
<td>ARIMA(3,0,1)(1,1,1)[12]</td>
<td>0.0630</td>
</tr>
<tr>
<td>ARIMA(4,0,3)(0,1,1)[12]</td>
<td>0.0648</td>
</tr>
<tr>
<td>ARIMA(3,0,3)(0,1,1)[12]</td>
<td>0.0639</td>
</tr>
<tr>
<td>ARIMA(4,0,2)(0,1,1)[12]</td>
<td>0.0648</td>
</tr>
<tr>
<td>ARIMA(3,0,2)(0,1,1)[12]</td>
<td>0.0644</td>
</tr>
<tr>
<td>ARIMA(2,1,3)(0,1,1)[12]</td>
<td>0.0634</td>
</tr>
<tr>
<td>ARIMA(2,1,4)(0,1,1)[12]</td>
<td>0.0632</td>
</tr>
<tr>
<td>ARIMA(2,1,5)(0,1,1)[12]</td>
<td>0.0640</td>
</tr>
</tbody>
</table>
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

```r
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")
```

Forecasts from ARIMA(3,0,1)(0,1,2)[12]
1. Backshift notation reviewed
2. Seasonal ARIMA models
3. ARIMA vs ETS
4. Lab session 12
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

<table>
<thead>
<tr>
<th>ETS model</th>
<th>ARIMA model</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETS(A,N,N)</td>
<td>ARIMA(0,1,1)</td>
<td>$\theta_1 = \alpha - 1$</td>
</tr>
<tr>
<td>ETS(A,A,N)</td>
<td>ARIMA(0,2,2)</td>
<td>$\theta_1 = \alpha + \beta - 2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\theta_2 = 1 - \alpha$</td>
</tr>
<tr>
<td>ETS(A,A,N)</td>
<td>ARIMA(1,1,2)</td>
<td>$\phi_1 = \phi$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\theta_1 = \alpha + \phi \beta - 1 - \phi$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\theta_2 = (1 - \alpha) \phi$</td>
</tr>
<tr>
<td>ETS(A,N,A)</td>
<td>ARIMA(0,0,m)(0,1,0)$_m$</td>
<td></td>
</tr>
<tr>
<td>ETS(A,A,A)</td>
<td>ARIMA(0,1,m + 1)(0,1,0)$_m$</td>
<td></td>
</tr>
<tr>
<td>ETS(A,A,A)</td>
<td>ARIMA(1,0,m + 1)(0,1,0)$_m$</td>
<td></td>
</tr>
</tbody>
</table>
Outline

1. Backshift notation reviewed
2. Seasonal ARIMA models
3. ARIMA vs ETS
4. Lab session 12
Lab Session 12