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 $B^{12}y_t = y_{t-12}.$
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:

$$B y_t = y_{t-1}.$$ 

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

- 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 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:**

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

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

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:**

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

\[
\begin{array}{c}
\uparrow \\
AR(1)
\end{array}
\begin{array}{c}
\uparrow \\
\text{First difference}
\end{array}
\begin{array}{c}
\uparrow \\
MA(1)
\end{array}
\]
Backshift notation for ARIMA

- **ARMA model:**

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

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

  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) \quad First \quad MA(1)

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

$$(1 - \phi_1B - \cdots - \phi_pB^p) (1 - B)^d y_t = c + (1 + \theta_1B + \cdots + \theta_qB^q)e_t$$

↑  ↑  ↑

AR($p$)  $d$ differences  MA($q$)
1 Backshift notation reviewed

2 Seasonal ARIMA models

3 ARIMA vs ETS

4 Lab session 12
### Seasonal ARIMA models

<table>
<thead>
<tr>
<th>ARIMA</th>
<th>( (p, d, q) )</th>
<th>( (P, D, Q)_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>↑ Non-seasonal part of the model</td>
<td>↑ Seasonal part of the model</td>
</tr>
</tbody>
</table>

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 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:

\[y_t = (1 + \phi_1)y_{t−1} − \phi_1 y_{t−2} + (1 + \Phi_1)y_{t−4} \]
\[− (1 + \phi_1 + \Phi_1 + \phi_1 \Phi_1)y_{t−5} + (\phi_1 + \phi_1 \Phi_1)y_{t−6} \]
\[− \Phi_1 y_{t−8} + (\Phi_1 + \phi_1 \Phi_1)y_{t−9} − \phi_1 \Phi_1 y_{t−10} \]
\[+ e_t + \theta_1 e_{t−1} + \Theta_1 e_{t−4} + \theta_1 \Theta_1 e_{t−5}.\]
In the US Census Bureau uses the following models most often:

<table>
<thead>
<tr>
<th>Model</th>
<th>Transformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA(0,1,1)(0,1,1)_m</td>
<td>with log tran.</td>
</tr>
<tr>
<td>ARIMA(0,1,2)(0,1,1)_m</td>
<td>with log tran.</td>
</tr>
<tr>
<td>ARIMA(2,1,0)(0,1,1)_m</td>
<td>with log tran.</td>
</tr>
<tr>
<td>ARIMA(0,2,2)(0,1,1)_m</td>
<td>with log tran.</td>
</tr>
<tr>
<td>ARIMA(2,1,2)(0,1,1)_m</td>
<td>with no tran.</td>
</tr>
</tbody>
</table>
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")
```
ggtsdisplay(diff(euretail, 4))
European quarterly retail trade

ggtsdisplay\((\text{diff}(\text{diff}(\text{euretail},4)))\)
- $d = 1$ 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))
ggtsdisplay(residuals(fit))
```

Forecasting using R  Seasonal ARIMA models  17
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
code
fit <- Arima(euretail, order=c(0,1,3), seasonal=c(0,1,1))
ggtsdisplay(residuals(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))
ggtsdisplay(residuals(fit))
```
European quarterly retail trade

residuals (fit)

Time

ACF

PACF

Forecasting using R

Seasonal ARIMA models
European quarterly retail trade

```r
res <- residuals(fit)
Box.test(res, lag=16, fitdf=4, type="Ljung")
```

```r
## Box-Ljung test
##
## data:  res
## X-squared = 7.0105, df = 12, p-value = 0.8569
```
European quarterly retail trade

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

Forecasts from ARIMA(0,1,3)(0,1,1)[4]
## Series: euretail
## ARIMA(1,1,2)(0,1,1)[4]
##
## Coefficients:
##    ar1    ma1    ma2    sma1
## 0.7345 -0.4655 0.2162 -0.8413
## s.e. 0.2239 0.1995 0.2096 0.1869
##
## sigma^2 estimated as 0.1592: log likelihood=-29.69
## AIC=69.37   AICc=70.51   BIC=79.76
auto.arima(euretail, stepwise=FALSE, approximation=FALSE)

## Series: euretail
## ARIMA(0,1,3)(0,1,1)[4]

## Coefficients:
## ma1    ma2    ma3    sma1
## 0.2625  0.3697  0.4194 -0.6615
## s.e.   0.1239  0.1260  0.1296  0.1555

## sigma^2 estimated as 0.1564: log likelihood=-28.7
## AIC=67.4   AICc=68.53   BIC=77.78
Corticosteroid drug sales

Forecasting using R
Seasonal ARIMA models
Corticosteroid drug sales

Seasonally differenced H02 scripts

Forecasting using R
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>
Corticosteroid 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
Corticosteroid drug sales

ggtsdisplay(residuals(fit))
Cortecosteroid drug sales

```r
Box.test(residuals(fit), lag=36, fitdf=6, type="Ljung")
```

```r
## Box-Ljung test
## data: residuals(fit)
## X-squared = 50.712, df = 30, p-value = 0.01045
```
fit <- auto.arima(h02, lambda=0, d=0, D=1, max.order=9, stepwise=FALSE, approximation=FALSE)
ggtsdisplay(residuals(fit))
Box.test(residuals(fit), lag=36, fitdf=8, type="Ljung")

##
## Box-Ljung test
##
## data: residuals(fit)
## X-squared = 44.766, df = 28, p-value = 0.02329
Cortecosteroid 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>

Forecasting using R

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