2.5 Seasonal ARIMA models
1. Seasonal ARIMA models
2. Lab session 17
3. ARIMA vs ETS
4. Lab session 18
Seasonal ARIMA models

**ARIMA** \((p, d, q)\)  \(\overset{\uparrow}{\text{Non-seasonal part of the model}}\)

\(\overset{\uparrow}{(P, D, Q)_m}\)  \(\text{Seasonal part of the model}\)

where \(m = \text{number of observations per year}\).
Seasonal ARIMA models

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

E.g., ARIMA\((1, 1, 1)(1, 1, 1)_4\) model (without constant)
\[(1-\phi_1B)(1-\Phi_1B^4)(1-B)(1-B^4)y_t = (1+\theta_1B)(1+\Theta_1B^4)\varepsilon_t.\]
Seasonal ARIMA models

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

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

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

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

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

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

- $\text{ARIMA}(0,1,1)(0,1,1)_m$ with log transformation
- $\text{ARIMA}(0,1,2)(0,1,1)_m$ with log transformation
- $\text{ARIMA}(2,1,0)(0,1,1)_m$ with log transformation
- $\text{ARIMA}(0,2,2)(0,1,1)_m$ with log transformation
- $\text{ARIMA}(2,1,2)(0,1,1)_m$ with no transformation
Understanding ARIMA models

<table>
<thead>
<tr>
<th>Long-term forecasts</th>
<th>$c = 0, d + D = 0$</th>
<th>$c = 0, d + D = 1$</th>
<th>$c = 0, d + D = 2$</th>
<th>$c = 0, d + D = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>zero</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-zero constant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>linear</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>quadratic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Forecast variance and $d + D$

- The higher the value of $d + D$, the more rapidly the prediction intervals increase in size.
- For $d + D = 0$, the long-term forecast standard deviation will go to the standard deviation of the historical data.
European quarterly retail trade

```
autoplot(euretail) +
  xlab("Year") + ylab("Retail index")
```
European quarterly retail trade

euretail %>% diff(lag=4) %>% autoplot()
European quarterly retail trade

euretail %>% diff(lag=4) %>% diff() %>% autoplot()
(fit <- auto.arima(euretail))

## Series: euretail
## ARIMA(1,1,2)(0,1,1)[4]
##
## Coefficients:
##       ar1    ma1    ma2    sma1
## 0.736 -0.466  0.216  -0.843
## s.e.  0.224  0.199  0.210  0.188
##
## sigma^2 estimated as 0.159:  log likelihood=-29.62
## AIC=69.24   AICc=70.38   BIC=79.63
(fit <- auto.arima(euretail, stepwise=TRUE, approximation=FALSE))

## Series: euretail
## ARIMA(1,1,2)(0,1,1)[4]
##
## Coefficients:
## ar1    ma1    ma2    sma1
## 0.736  -0.466  0.216  -0.843
## s.e.   0.224  0.199  0.210  0.188

## sigma^2 estimated as 0.159: log likelihood=-29.62
## AIC=69.24   AICc=70.38   BIC=79.63
European quarterly retail trade

```r
checkresiduals(fit, test=FALSE)
```

Residuals from ARIMA(1,1,2)(0,1,1)[4]

ACF

Lag

Residuals

Count

13
checkresiduals(fit, plot=FALSE)

##
## Ljung-Box test
##
## data: Residuals from ARIMA(1,1,2)(0,1,1)[4]
## Q* = 4.9, df = 4, p-value = 0.3
##
## Model df: 4. Total lags used: 8
**European quarterly retail trade**

```
forecast(fit, h=36) %>% autoplot()
```

Forecasts from ARIMA(1,1,2)(0,1,1)[4]
Corticosteroid drug sales

<table>
<thead>
<tr>
<th>Year</th>
<th>H02 sales (million scripts)</th>
<th>Log H02 sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>0.50</td>
<td>-0.8</td>
</tr>
<tr>
<td>2000</td>
<td>0.75</td>
<td>-0.4</td>
</tr>
<tr>
<td>2005</td>
<td>1.00</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Year
Cortecosteroid drug sales

```r
autoplot(diff(log(h02), 12), xlab="Year", main="Seasonally differenced H02 scripts")
```
Corticosteroid drug sales

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

```r
## Series: h02
## ARIMA(4,1,1)(2,1,2)[12]
## Box Cox transformation: lambda= 0
##
## Coefficients:
##           ar1   ar2   ar3   ar4  ma1   sar1
## sar2 -0.383 -1.202  0.496
## s.e.  0.118  0.249  0.214
```

18
Corticosteroid drug sales

checkresiduals(fit)

Residuals from ARIMA(4,1,1)(2,1,2)[12]

Lag
ACF
0
10
20
30
−0.2 −0.1 0.0 0.1 0.2
residuals
count

## Ljung-Box test

data: Residuals from ARIMA(4,1,1)(2,1,2)[12]

Q* = 16, df = 15, p-value = 0.4

Model df: 9. Total lags used: 24
Cortecosteroid drug sales

## Ljung-Box test

## data: Residuals from ARIMA(4,1,1)(2,1,2)[12]
## Q* = 16, df = 15, p-value = 0.4

## Model df: 9. Total lags used: 24
Cortecosteroid drug sales

Training data: July 1991 to June 2006

Test data: July 2006–June 2008

def getrmse(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)
Cortecosteroid drug sales

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA(4,1,2)(2,1,2)[12]</td>
<td>0.0614</td>
</tr>
<tr>
<td>ARIMA(4,1,1)(2,1,2)[12]</td>
<td>0.0615</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(2,1,4)(0,1,1)[12]</td>
<td>0.0632</td>
</tr>
<tr>
<td>ARIMA(2,1,3)(0,1,1)[12]</td>
<td>0.0634</td>
</tr>
<tr>
<td>ARIMA(4,1,2)(1,1,2)[12]</td>
<td>0.0634</td>
</tr>
<tr>
<td>ARIMA(3,1,2)(2,1,2)[12]</td>
<td>0.0636</td>
</tr>
<tr>
<td>ARIMA(3,0,3)(0,1,1)[12]</td>
<td>0.0639</td>
</tr>
<tr>
<td>ARIMA(2,1,5)(0,1,1)[12]</td>
<td>0.0640</td>
</tr>
<tr>
<td>ARIMA(3,0,1)(0,1,1)[12]</td>
<td>0.0644</td>
</tr>
<tr>
<td>ARIMA(3,0,2)(0,1,1)[12]</td>
<td>0.0644</td>
</tr>
<tr>
<td>ARIMA(3,0,2)(2,1,0)[12]</td>
<td>0.0645</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.
- Use the best model available, even if it does not pass all tests.
Cortecosteroid drug sales

```r
fit <- Arima(h02, order=c(4,1,1), seasonal=c(2,1,2), lambda=0)
autoplot(forecast(fit)) + xlab("Year") + ylab("H02 sales (million scripts)") + ylim(0.3,1.8)
```

Forecasts from ARIMA(4,1,1)(2,1,2)[12]
Cortecosteroid drug sales

```r
fit <- Arima(h02, order=c(4,1,2), seasonal=c(2,1,2), lambda=0)
autoplot(forecast(fit)) + xlab("Year") +
ylab("H02 sales (million scripts)") + ylim(0.3,1.8)
```

Forecasts from ARIMA(4,1,2)(2,1,2)[12]
Cortecosteroid drug sales

```r
fit <- Arima(h02, order=c(3,0,1), seasonal=c(0,1,2), lambda=0)
autoplot(forecast(fit)) + xlab("Year") +
  ylab("H02 sales (million scripts)") + ylim(0.3,1.8)
```

Forecasts from ARIMA(3,0,1)(0,1,2)[12]
1. Seasonal ARIMA models

2. Lab session 17

3. ARIMA vs ETS

4. Lab session 18
Lab Session 17
1. Seasonal ARIMA models
2. Lab session 17
3. ARIMA vs ETS
4. Lab session 18
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>
</tbody>
</table>
| ETS(A,A,N) | ARIMA(0,2,2) | $\theta_1 = \alpha + \beta - 2$  
$\theta_2 = 1 - \alpha$ |
| ETS(A,Ad,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,Ad,A) | ARIMA(1,0,m + 1)(0,1,0)$_m$ |  |
Outline

1. Seasonal ARIMA models
2. Lab session 17
3. ARIMA vs ETS
4. Lab session 18
Lab Session 18