MONASH University

MONASH BUSINESS SCHOOL

Time Series and Forecasting

ARIMA models

Outline

1 Stationarity
2 Differencing
3 Non-seasonal ARIMA models
4 Estimation and order selection
5 ARIMA modelling in R
6 Seasonal ARIMA models
7 ARIMA vs ETS

Stationarity

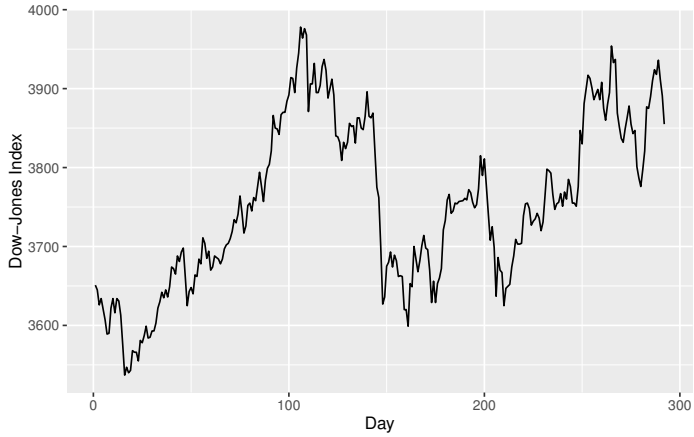
Definition

If $\{y_t\}$ is a stationary time series, then for all s, the distribution of $(y_t, \ldots, y_{t+s})$ does not depend on t.

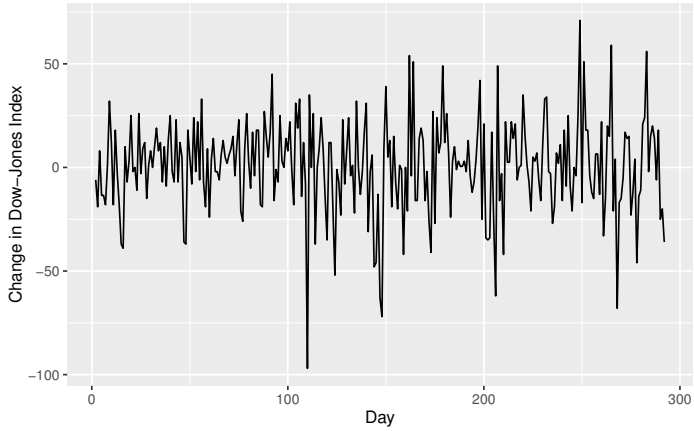
A **stationary series** is:

- roughly horizontal
- constant variance
- no patterns predictable in the long-term

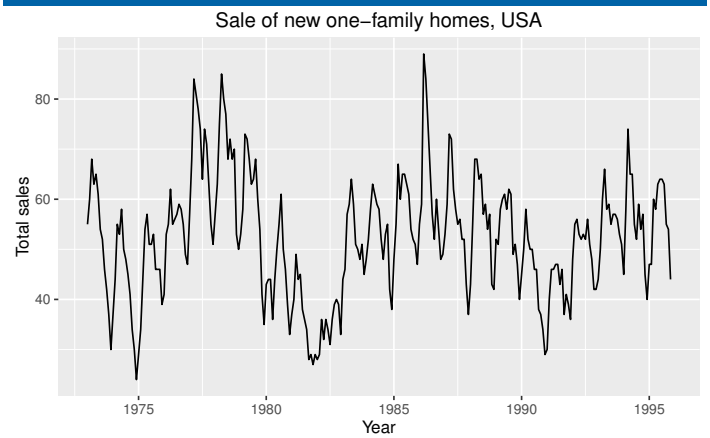
Stationary?



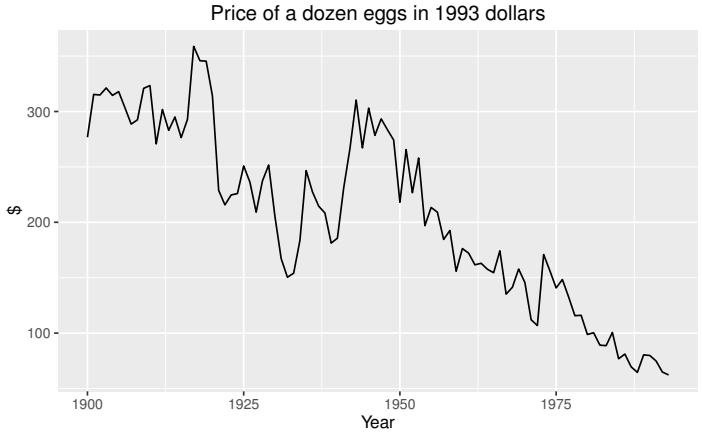
Stationary?



Stationary?

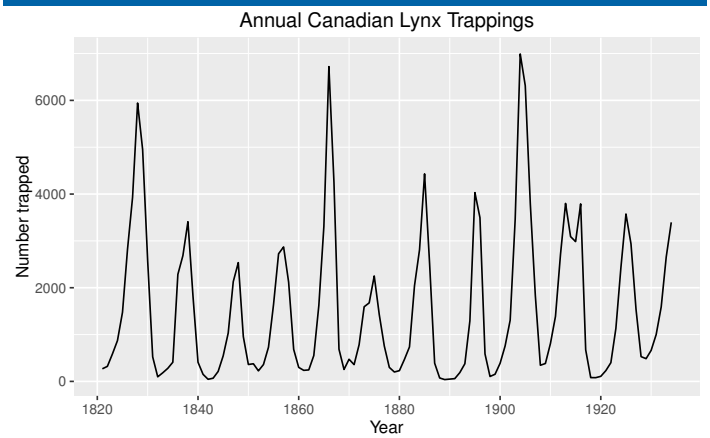


Stationary?



Stationary?

Stationary?



Stationarity

Definition

If $\{y_t\}$ is a stationary time series, then for all s, the distribution of $(y_t, \ldots, y_{t+s})$ does not depend on t.

Transformations help to **stabilize the variance**.

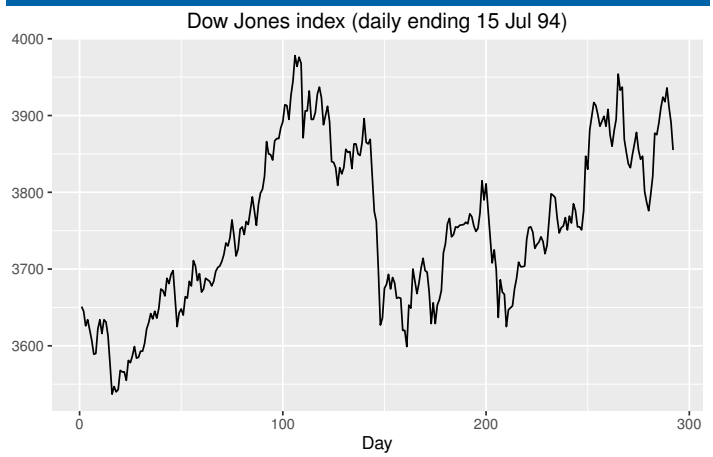
For ARIMA modelling, we also need to **stabilize the mean**.

Non-stationarity in the mean

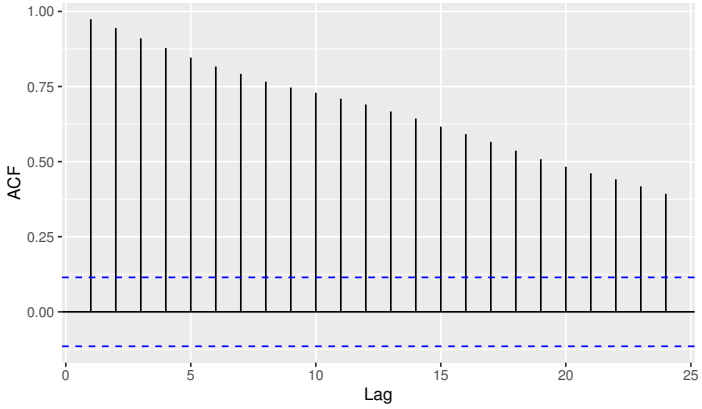
Identifying non-stationary series

- time plot.
- The ACF of stationary data drops to zero relatively quickly
- The ACF of non-stationary data decreases slowly.
- For non-stationary data, the value of r_1 is often large and positive.

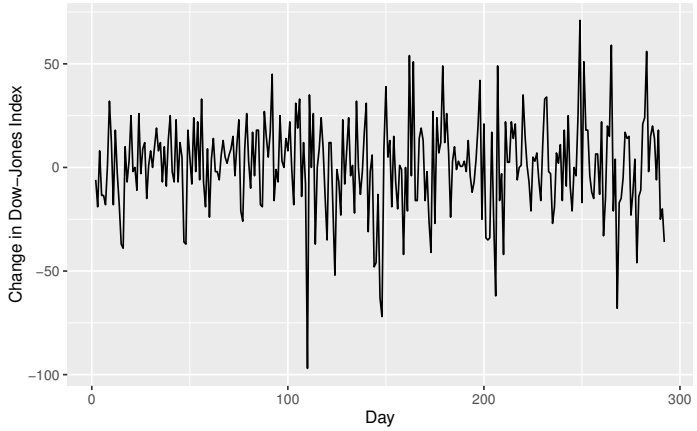
Example: Dow-Jones index



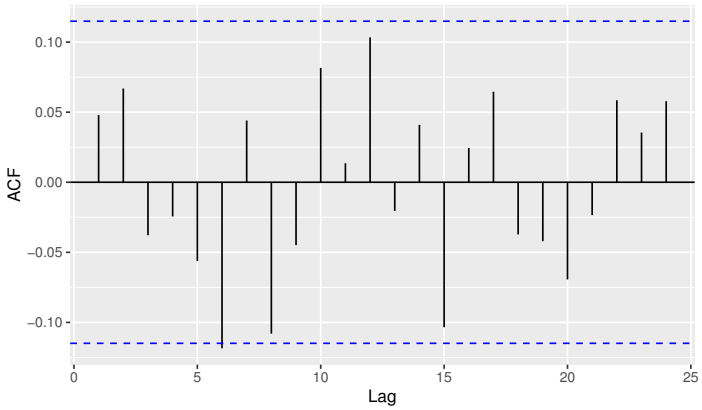
Example: Dow-Jones index



Example: Dow-Jones index



Example: Dow-Jones index



Outline

1 Stationarity
2 Differencing
3 Non-seasonal ARIMA models
4 Estimation and order selection
5 ARIMA modelling in R
6 Seasonal ARIMA models
7 ARIMA vs ETS

Differencing

- Differencing helps to **stabilize the mean**.
- The differenced series is the *change* between each observation in the original series: $y'_t = y_t - y_{t-1}$.
- The differenced series will have only $T - 1$ values since it is not possible to calculate a difference y'_1 for the first observation.

Second-order differencing

Occasionally the differenced data will not appear stationary and it may be necessary to difference the data a second time:

$$\begin{aligned} y''_t &= y'_t - y'_{t-1} \\ &= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \\ &= y_t - 2y_{t-1} + y_{t-2}. \end{aligned}$$

- y''_t will have $T - 2$ values.
- In practice, it is almost never necessary to go beyond second-order differences.

Seasonal differencing

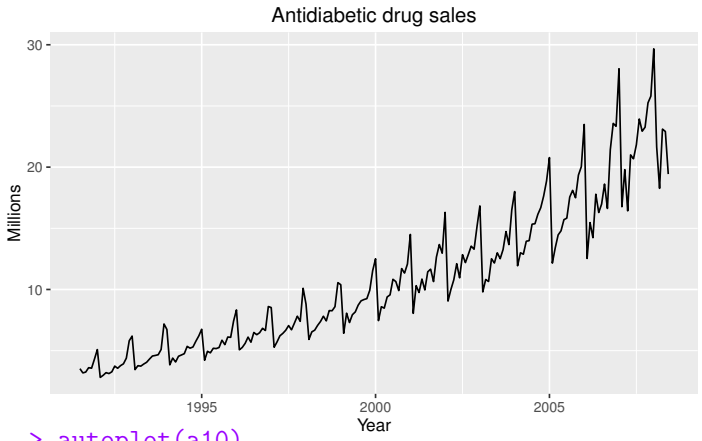
A seasonal difference is the difference between an observation and the corresponding observation from the previous year.

$$y'_t = y_t - y_{t-m}$$

where m = number of seasons.

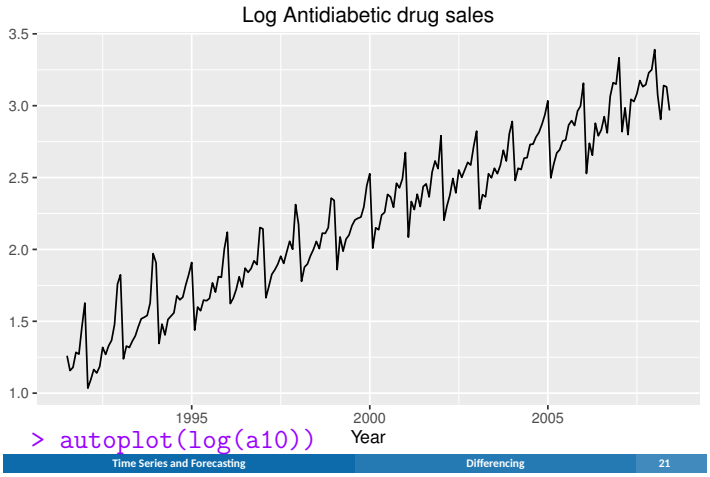
- For monthly data $m = 12$.
- For quarterly data $m = 4$.

Antidiabetic drug sales



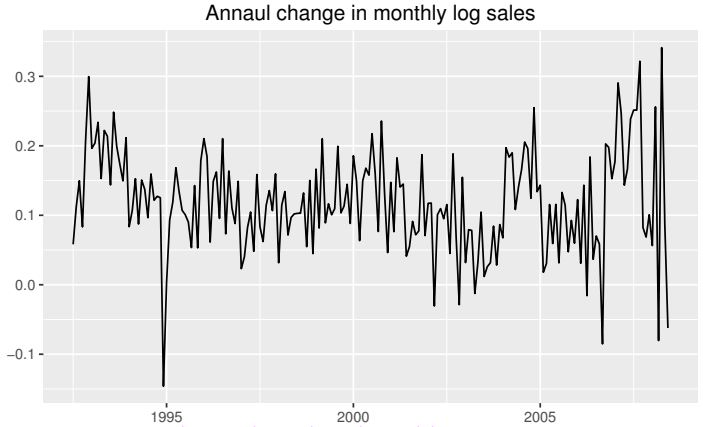

```
> autoplot(a10)
```


Antidiabetic drug sales



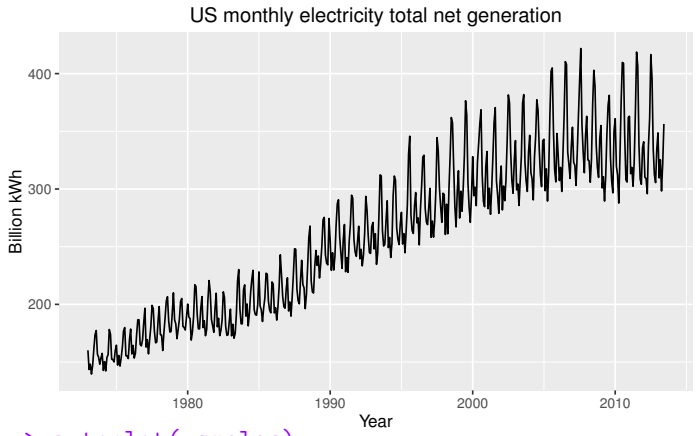

```
> autoplot(log(a10))
```


Antidiabetic drug sales



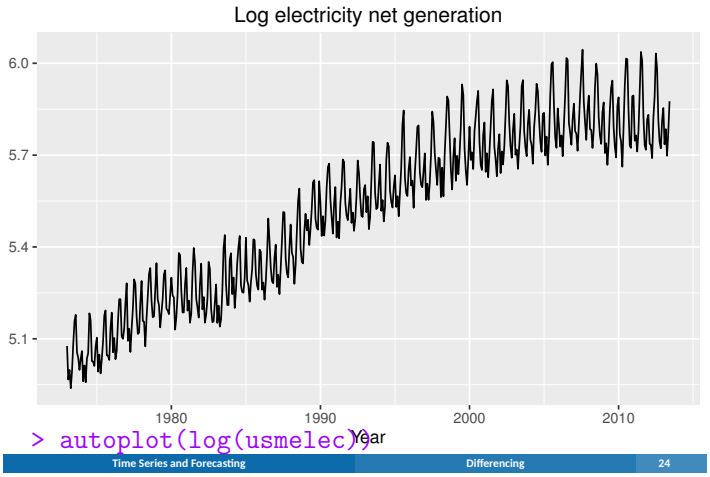

```
> autoplot(diff(log(a10),12))
```


Electricity production



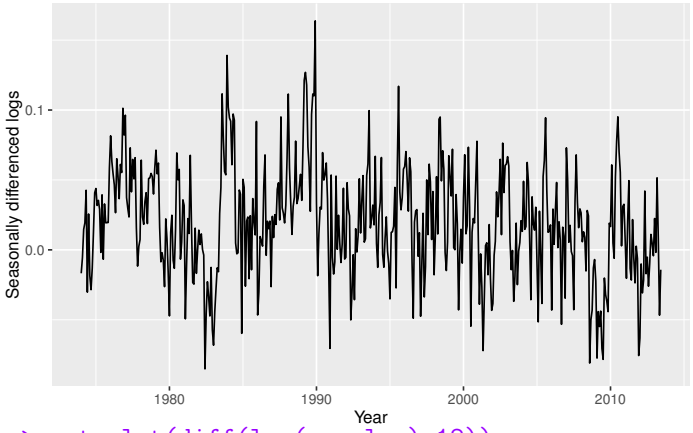

```
> autoplot(usmelec)
```


Electricity production



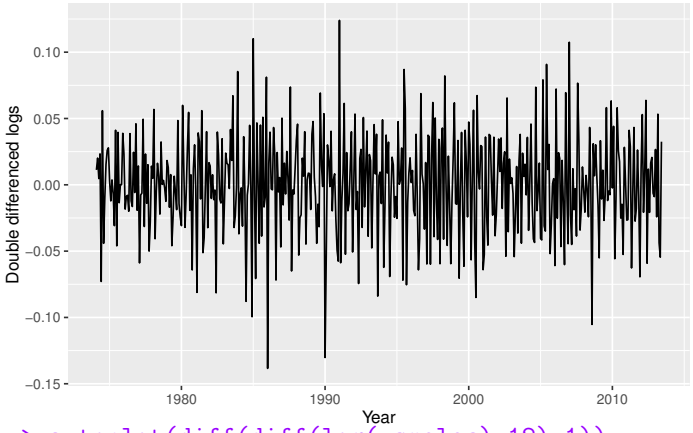

```
> autoplot(log(usmelec))
```

Electricity production




```
> autoplot(diff(log(usmelec),12))
```


Electricity production




```
> autoplot(diff(diff(log(usmelec),12),1))
```


Electricity production

- Seasonally differenced series is closer to being stationary.
- Remaining non-stationarity can be removed with further first difference.

If $y'_t = y_t - y_{t-12}$ denotes seasonally differenced series, then twice-differenced series is

$$\begin{aligned} y^*_t &= y'_t - y'_{t-1} \\ &= (y_t - y_{t-12}) - (y_{t-1} - y_{t-13}) \\ &= y_t - y_{t-1} - y_{t-12} + y_{t-13}\,. \end{aligned}$$

Seasonal differencing

When both seasonal and first differences are applied…

- it makes no difference which is done first—the result will be the same.
- If seasonality is strong, we recommend that seasonal differencing be done first because sometimes the resulting series will be stationary and there will be no need for further first difference.

It is important that if differencing is used, the differences are interpretable.

Outline

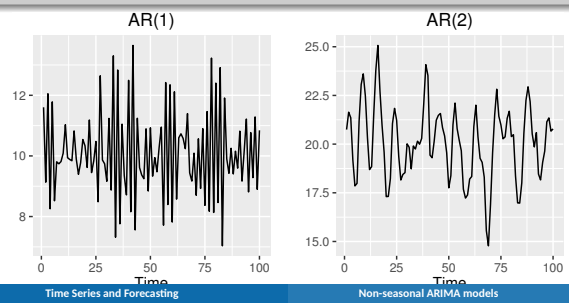
1 Stationarity
2 Differencing
3 Non-seasonal ARIMA models
4 Estimation and order selection
5 ARIMA modelling in R
6 Seasonal ARIMA models
7 ARIMA vs ETS

Autoregressive models

Autoregressive (AR) models:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t,$$

where e_t is white noise.This is a multiple regression with **lagged values** of y_t as predictors.



AR(1) model

$$y_t = c + \phi_1 y_{t-1} + e_t$$

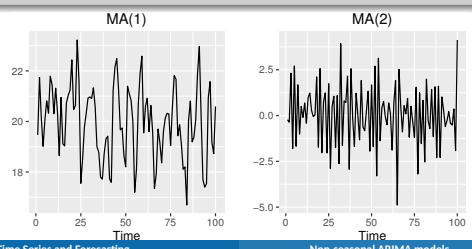
- When $\phi_1 = 0$, y_t is **equivalent to WN**
- When $\phi_1 = 1$ and $c = 0$, y_t is **equivalent to a RW**
- When $\phi_1 = 1$ and $c \neq 0$, y_t is **equivalent to a RW with drift**
- When $\phi_1 < 0$, y_t tends to **oscillate between positive and negative values**.

Moving Average (MA) models

Moving Average (MA) models:

$$y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q},$$

where e_t is white noise.This is a multiple regression with **past *errors*** as predictors. *Don't confuse this with moving average smoothing!*



ARIMA models

Autoregressive Moving Average models:

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

- Predictors include both **lagged values of y_t and lagged errors.**
- Some constraints on coefficients to ensure identifiability and numerical stability.

Autoregressive Integrated Moving Average models

- Combine ARMA model with **differencing**.
- $(1-B)^d y_t$ follows an ARMA model.

ARIMA models

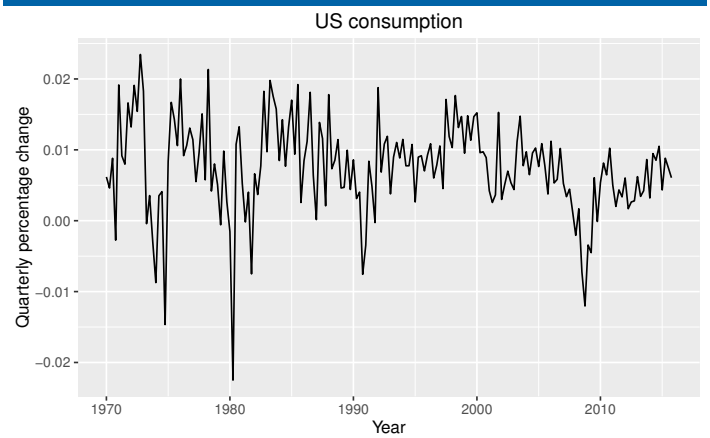
Autoregressive Integrated Moving Average models

ARIMA(p, d, q) model

AR: p = order of the autoregressive part
I: d = degree of first differencing involved
MA: q = order of the moving average part.

- White noise model: ARIMA(0,0,0)
- Random walk: ARIMA(0,1,0) with no constant
- Random walk with drift: ARIMA(0,1,0) with const.
- AR(p): ARIMA(p,0,0)
- MA(q): ARIMA(0,0,q)

US personal consumption



US personal consumption

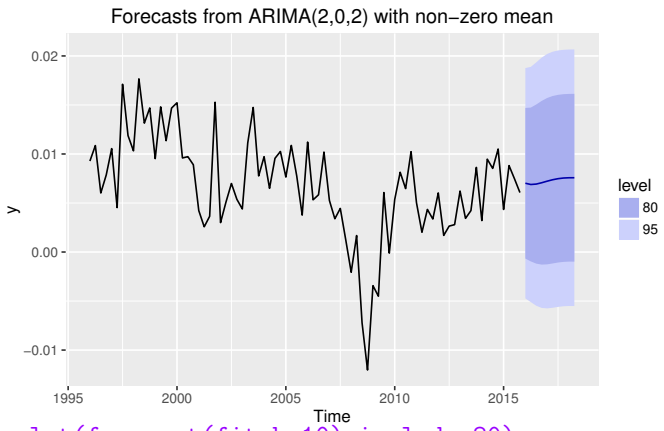
```
> fit <- auto.arima(usconsumption[, 1])
> fit
Series: usconsumption[, 1]
ARIMA(2,0,2) with non-zero mean

Coefficients:
         ar1      ar2      ma1     ma2  intercept
      1.3858  -0.5787  -1.1797  0.5633     0.0075
s.e.  0.2470   0.2040   0.2305  0.1409     0.0009

sigma^2 estimated as 3.593e-05:  log likelihood=682.77
AIC=-1353.54   AICc=-1353.06   BIC=-1334.25
```

ARIMA(2,0,2) model:

$$y_t = c + 1.3858 y_{t-1} - 0.578 y_{t-2} + e_t - 1.17974 e_{t-1} + 0.5633 e_{t-2},$$

where e_t is white noise with std dev $0.0060 = \sqrt{3.593 \times 10^{-5}}$.

US personal consumption




```
plot(forecast(fit,h=10),include=80)
```


Understanding ARIMA models

- If $c = 0$ and $d = 0$, the long-term forecasts will go to zero.
- If $c = 0$ and $d = 1$, the long-term forecasts will go to a non-zero constant.
- If $c = 0$ and $d = 2$, the long-term forecasts will follow a straight line.
- If $c \neq 0$ and $d = 0$, the long-term forecasts will go to the mean of the data.
- If $c \neq 0$ and $d = 1$, the long-term forecasts will follow a straight line.
- If $c \neq 0$ and $d = 2$, the long-term forecasts will follow a quadratic trend.

Understanding ARIMA models

Forecast variance and d

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

Cyclic behaviour

- For cyclic forecasts, $p > 2$ and some restrictions on coefficients are required.
- If $p = 2$, we need $\phi_1^2 + 4\phi_2 < 0$. Then average cycle of length

$$(2\pi)/\left[\arccos(-\phi_1(1-\phi_2)/(4\phi_2))\right].$$

Outline

1 Stationarity
2 Differencing
3 Non-seasonal ARIMA models
4 Estimation and order selection
5 ARIMA modelling in R
6 Seasonal ARIMA models
7 ARIMA vs ETS

Maximum likelihood estimation

Having identified the model order, we need to estimate the parameters $c, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$.

- MLE is very similar to least squares estimation obtained by minimizing

$$\sum_{t-1}^{T} e_t^2.$$

- The `Arima()` command allows CLS or MLE estimation.
- Non-linear optimization must be used in either case.
- Different software will give different estimates.

Information criteria

Akaike's Information Criterion (AIC):

$$\text{AIC} = -2\log(L) + 2(p+q+k+1),$$

where L is the likelihood of the data, $k = 1$ if $c \neq 0$ and $k = 0$ if $c = 0$.

Corrected AIC:

$$\text{AIC}_c = \text{AIC} + \frac{2(p+q+k+1)(p+q+k+2)}{T-p-q-k-2}.$$

Bayesian Information Criterion:

$$\text{BIC} = \text{AIC} + (\log(T) - 2)(p+q+k-1).$$

Good models are obtained by minimizing either the AIC, AIC_c or BIC. Our preference is to use the AIC_c.

Outline

1 Stationarity
2 Differencing
3 Non-seasonal ARIMA models
4 Estimation and order selection
5 ARIMA modelling in R
6 Seasonal ARIMA models
7 ARIMA vs ETS

How does auto.arima() work?

A non-seasonal ARIMA(p, d, q) process

Need to select appropriate orders: p, q, d and whether a constant is required.

Hyndman and Khandakar (JSS, 2008) algorithm:

- Select no. differences d and D via unit root tests.
- Select p, q and c by minimising AICc.
- Use stepwise search to traverse model space.

How does auto.arima() work?

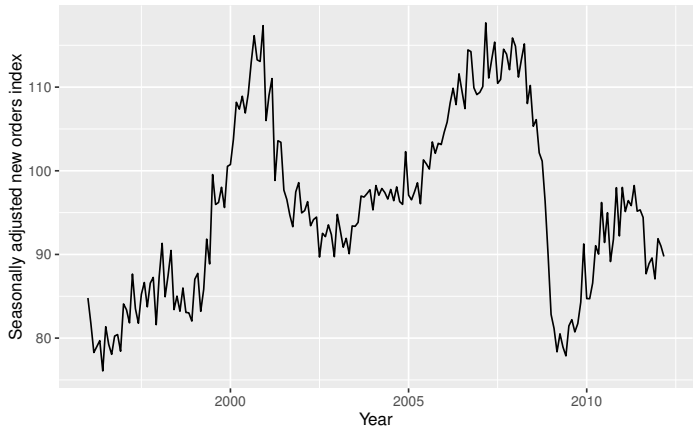
Step 1: Select current model (with smallest AICc) from:
ARIMA($2, d, 2$)
ARIMA($0, d, 0$)
ARIMA($1, d, 0$)
ARIMA($0, d, 1$)

Step 2: Consider variations of current model:
- vary one of p, q from current model by ± 1
- p, q both vary from current model by ± 1
- Include/exclude c from current model

Model with lowest AICc becomes current model.

Repeat Step 2 until no lower AICc can be found.

Seasonally adjusted electrical equipment



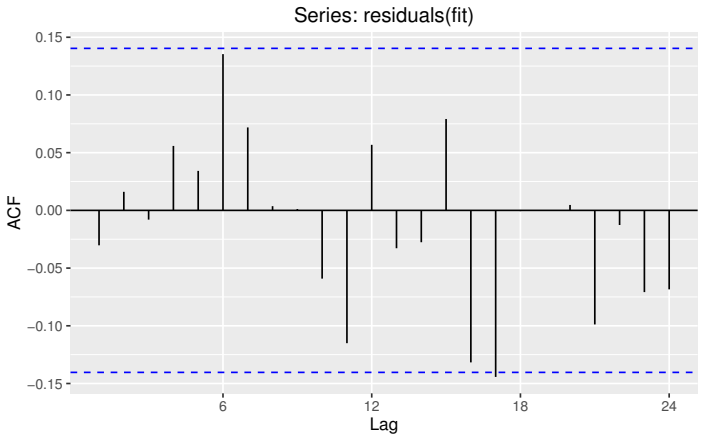
Seasonally adjusted electrical equipment

- Time plot shows sudden changes, particularly big drop in 2008/2009 due to global economic environment. Otherwise nothing unusual and no need for data adjustments.
- No evidence of changing variance, so no Box-Cox transformation.
- Data are clearly non-stationary, suggesting first differences.

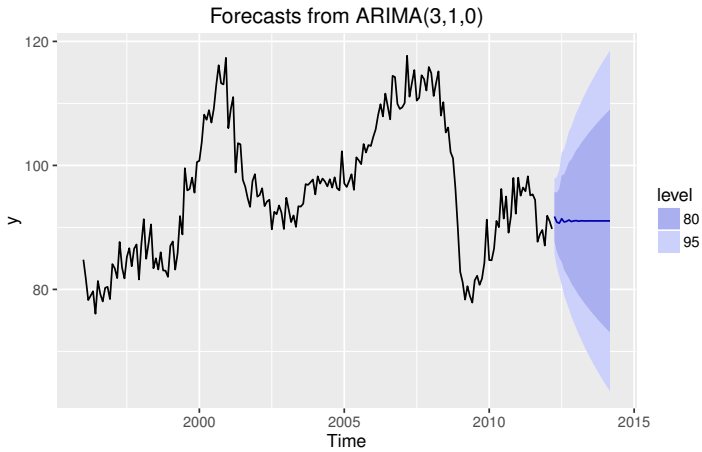
Seasonally adjusted electrical equipment

```
> fit <- auto.arima(eeadj)
> summary(fit)
ARIMA(3,1,0)

Coefficients:
          ar1      ar2     ar3
      -0.3418  -0.0426  0.3185
s.e.   0.0681   0.0725  0.0682

sigma^2 estimated as 9.639:  log likelihood=-493.8
AIC=995.6   AICc=995.81   BIC=1008.67

> ggAcf(residuals(fit))

> autoplot(forecast(fit))
```

Seasonally adjusted electrical equipment



Seasonally adjusted electrical equipment



Prediction intervals

- Prediction intervals **increase in size with forecast horizon.**
- Prediction intervals can be difficult to calculate by hand
- Calculations assume residuals are **uncorrelated** and **normally distributed**.
- Prediction intervals tend to be too narrow.
  - the uncertainty in the parameter estimates has not been accounted for.
  - the ARIMA model assumes historical patterns will not change during the forecast period.
  - the ARIMA model assumes uncorrelated future errors

Outline

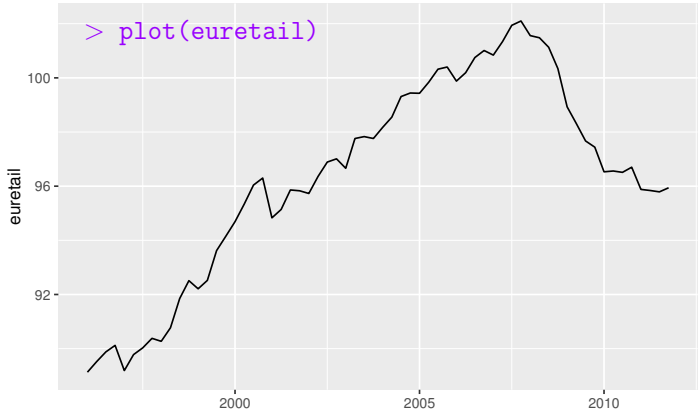
1 Stationarity
2 Differencing
3 Non-seasonal ARIMA models
4 Estimation and order selection
5 ARIMA modelling in R
6 **Seasonal ARIMA models**
7 ARIMA vs ETS

Seasonal ARIMA models

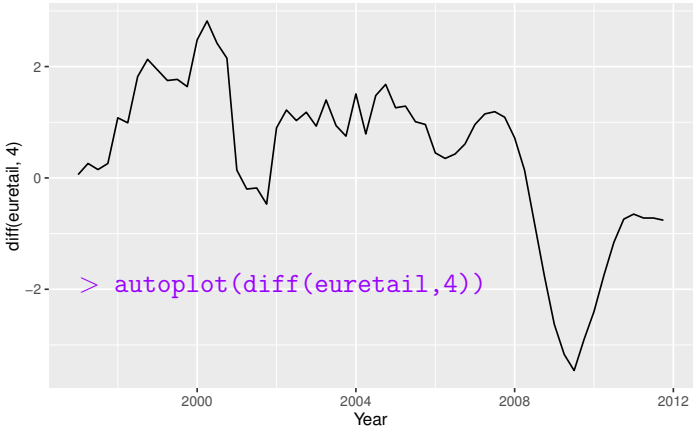
$$\text{ARIMA}\ \underbrace{(p,d,q)}_{\substack{\uparrow \\ \text{Non-seasonal part of the model}}}\ \underbrace{(P,D,Q)_m}_{\substack{\uparrow \\ \text{Seasonal part of the model}}}$$

where m = number of periods per season.

European quarterly retail trade

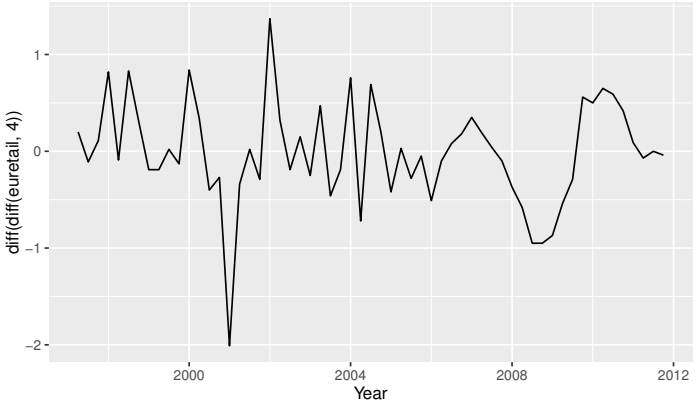


European quarterly retail trade




```
> autoplot(diff(euretail,4))
```


European quarterly retail trade




```
> autoplot(diff(diff(euretail,4)))
```

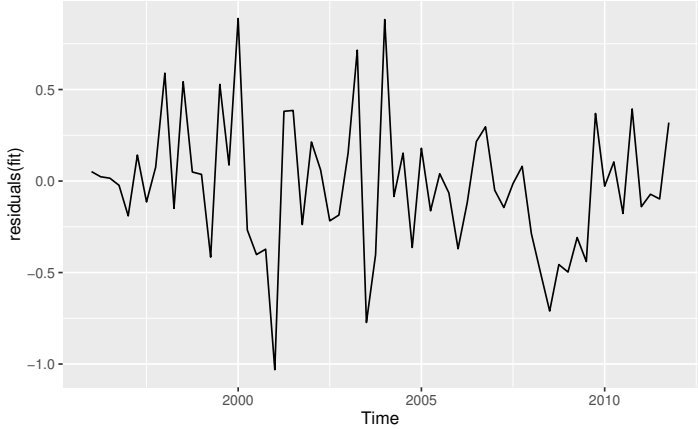

European quarterly retail trade

```
> 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
```

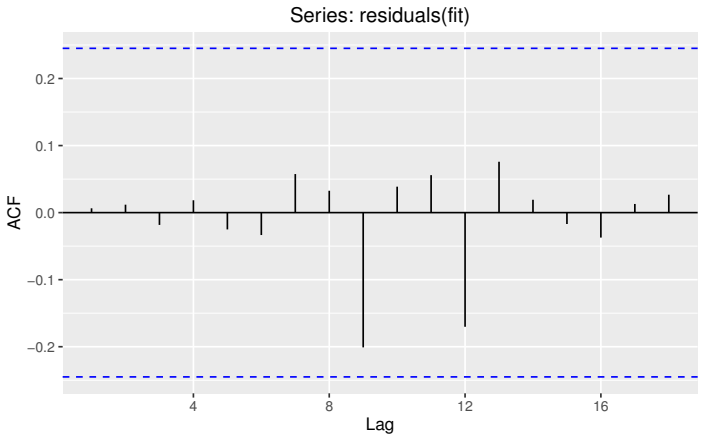

European quarterly retail trade

```
> auto.arima(euretail, stepwise=FALSE,
    approximation=FALSE)
ARIMA(0,1,3)(0,1,1)[4]

Coefficients:
        ma1    ma2    ma3    sma1
      0.263  0.369  0.420  -0.664
s.e.  0.124  0.126  0.129   0.155

sigma^2 estimated as 0.156:  log likelihood=-28.63
AIC=67.26   AICc=68.39   BIC=77.65
```

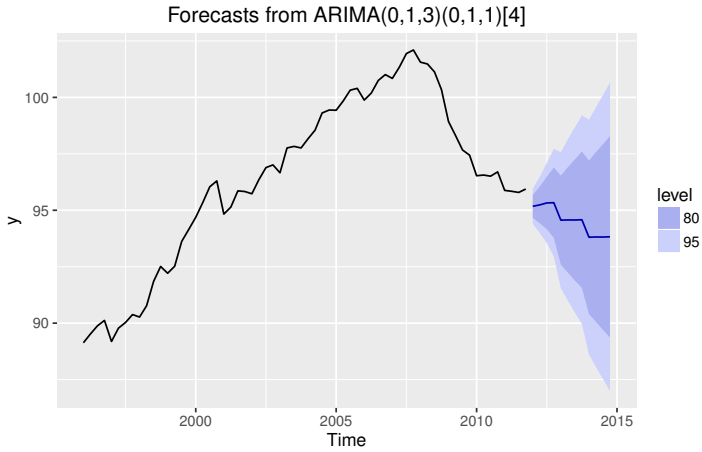

European quarterly retail trade



European quarterly retail trade



European quarterly retail trade



Outline

1 Stationarity

2 Differencing

3 Non-seasonal ARIMA models

4 Estimation and order selection

5 ARIMA modelling in R

6 Seasonal ARIMA models

7 ARIMA vs ETS

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

Simple exponential smoothing

- Forecasts equivalent to **ARIMA(0,1,1)**.
- Parameters: $\theta_1 = \alpha - 1$.

Holt's method

- Forecasts equivalent to **ARIMA(0,2,2)**.
- Parameters: $\theta_1 = \alpha + \beta - 2$ and $\theta_2 = 1 - \alpha$.

Damped Holt's method

- Forecasts equivalent to **ARIMA(1,1,2)**.
- Parameters: $\phi_1 = \phi$, $\theta_1 = \alpha + \phi\beta - 2$, $\theta_2 = (1 - \alpha)\phi$.

Holt-Winters' additive method

- Forecasts equivalent to **ARIMA(0,1,m+1)(0,1,0)$_m$**.
- Parameter restrictions because ARIMA has $m + 1$ parameters whereas HW uses only three parameters.

Holt-Winters' multiplicative method

- No ARIMA equivalence