ARMA

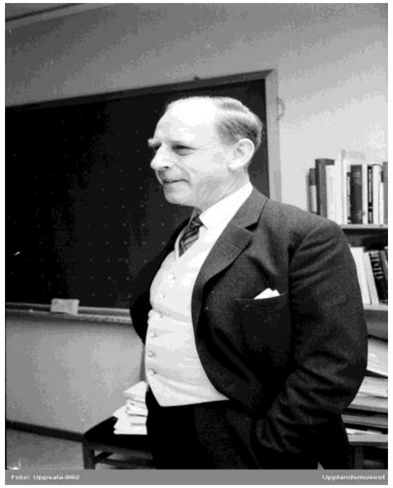
Simulation and Fitting

Stationary Time Series: ARMA

- Wold Decomposition: Herman Wold (1908-1992) proved that any time series can be represented as a linear combination of white noise:

$$y_t = W_t + a_1 W_{t-1} + a_2 W_{t-2} + \cdots$$

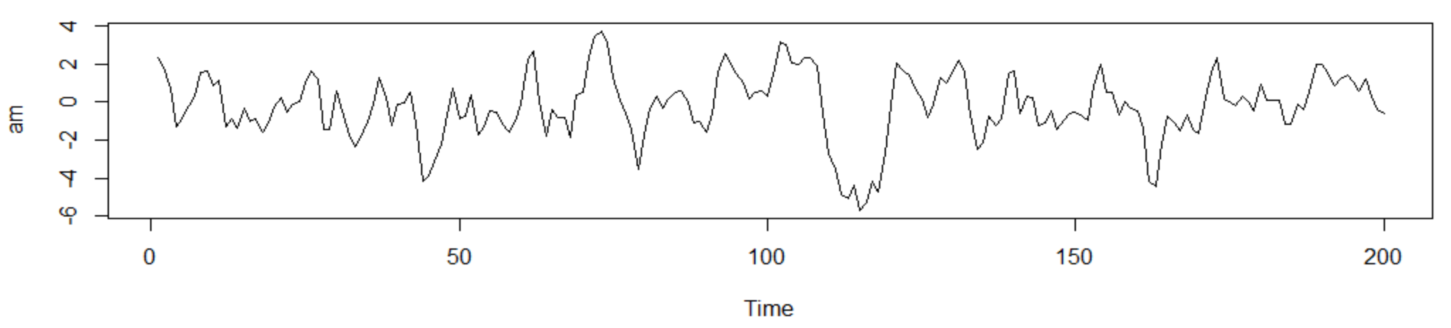
where $a_1, a_2, \ldots$ are constants



Simulating ARMA in R

- ARMA model can be simulated using function arima.sim() by specifying the order option as c(p,d=0,q) in the list form, where
  - p = Order of AR
  - d = Order of Differencing
  - q = Order of MA

```
am <- arima.sim(list(order=c(p=1,d=0,q=1),ar=0.7,ma=0.3),n=200)
plot(am)
```



Identifying AR and MA Models

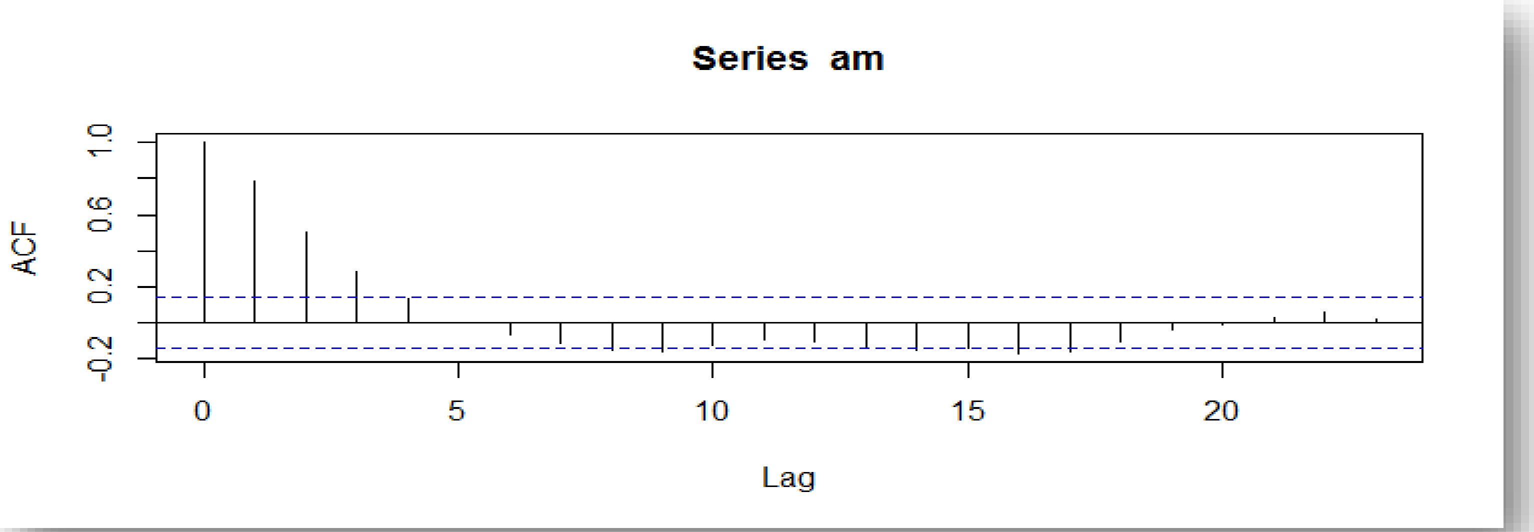
- It is very hard to identify AR and MA models with simple graphs like line graph because many times they look similar.
- We can identify them with the help of graphs ACF (Auto-correlation function) and PACF (Partial Auto-Correlation Function)

Auto-Correlation

- Let x_t denote the set of values of a time series at time t and x_{t-h} denotes the set of values of a time series at time $t-h,$ in other words h time periods apart
- Auto-Correlation with h time period apart is defined as

$$\frac{Covariance(x_t, x_{t-h})}{Std\ Dev(x_t) Std\ Dev(x_{t-h})}$$

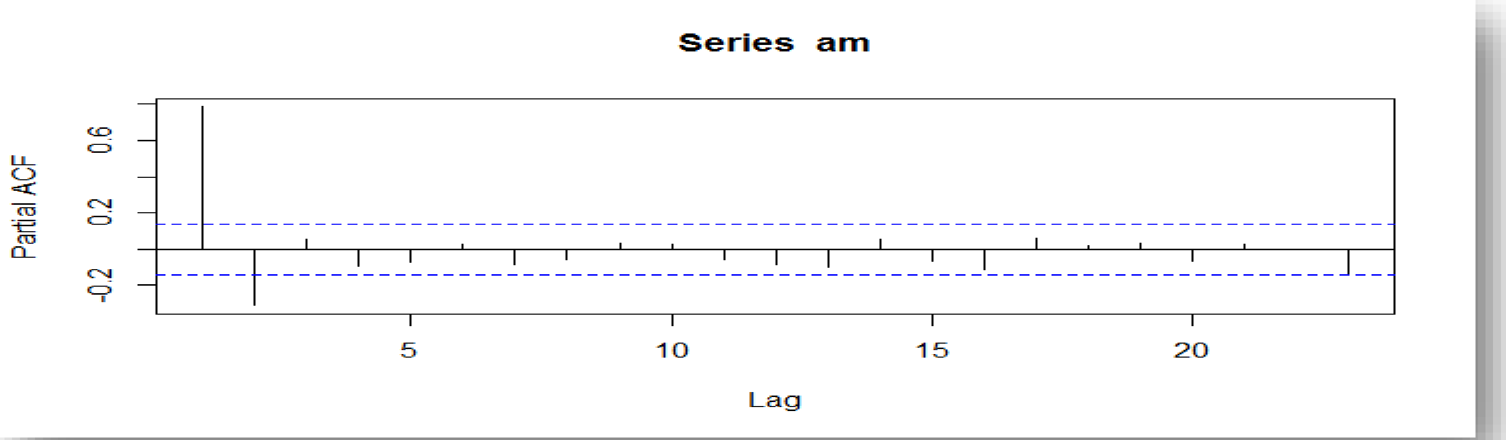
- The graph of different possible h lags is plotted to form ACF graph



Partial Auto-Correlation

- Partial correlation is a conditional correlation
- It is the correlation between two variables under the assumption that we know and take into account the values of some other set of variables
- Partial Auto-correlation of a stationary process x_t, is calculated as $\emptyset_{11} = corr(x_1, x_0)$ & $\emptyset_{hh} = corr(x_h - \widehat{x_h}, x_0 - \widehat{x_0}), h \geq 2$

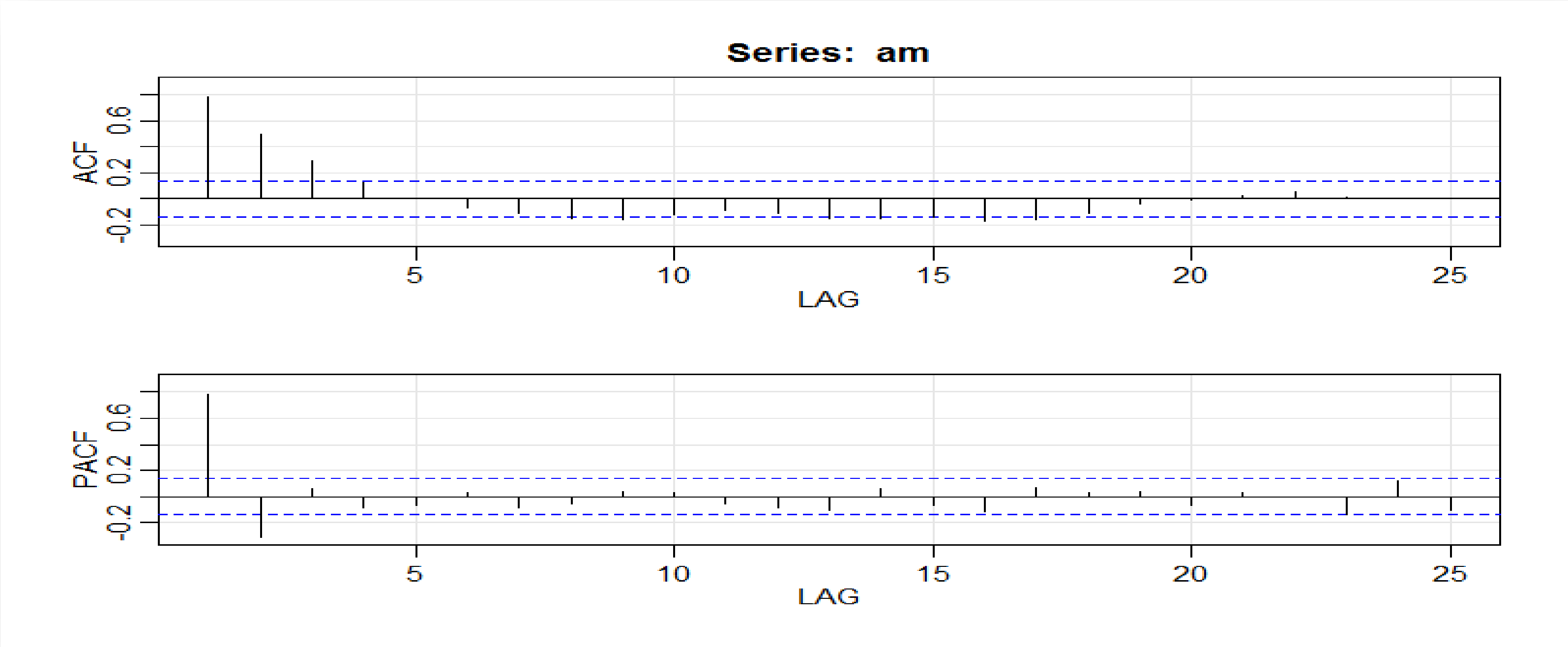
  where $\widehat{x_h}$ is the regression of x_h on $\{x_1, x_2, \dots x_{h-1}\}$ and $\widehat{x_0}$ is the regression of x_0 on $\{x_1, x_2, \dots x_{h-1}\}$
- The graph of different possible h lags is plotted to form PACF graph



Function acf2()

- In package **astsa**, there is a function *acf2()* which display ACF as well as PACF graphs at one go.

```
library(astsa)
acf2(am)
```



Function sarima()

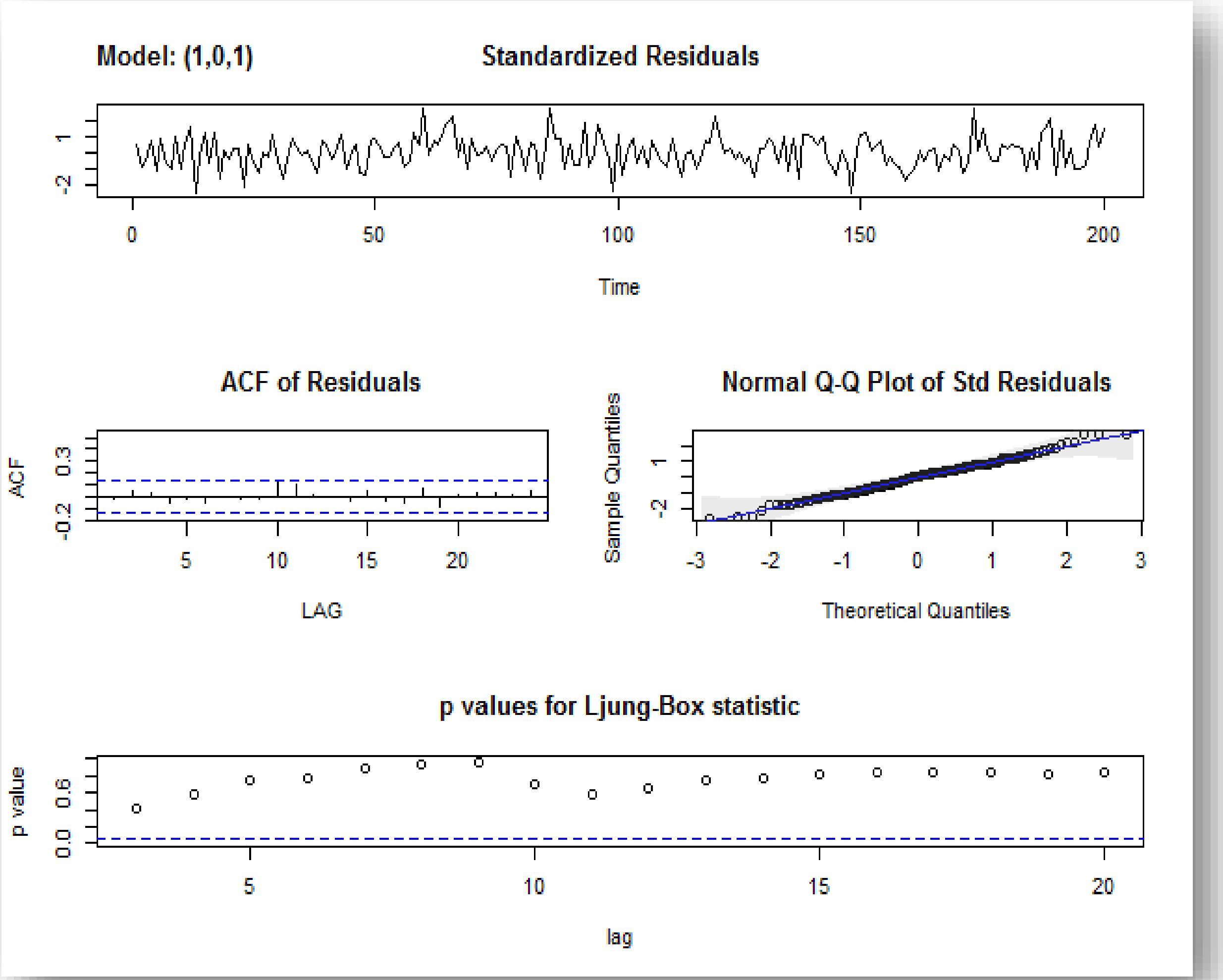
- Function sarima() from package astsa fits ARIMA as well as seasonal ARIMA models

Syntax : sarima(x, p, d ,q, P, D, Q, S, …)

Where

x : univariate time series
p : AR order
d : difference order
q : MA order
P : SAR order; use only for seasonal models
D : seasonal difference; use only for seasonal models
Q : SMA order; use only for seasonal models
S : seasonal period; use only for seasonal models

Output Graphs of sarima()



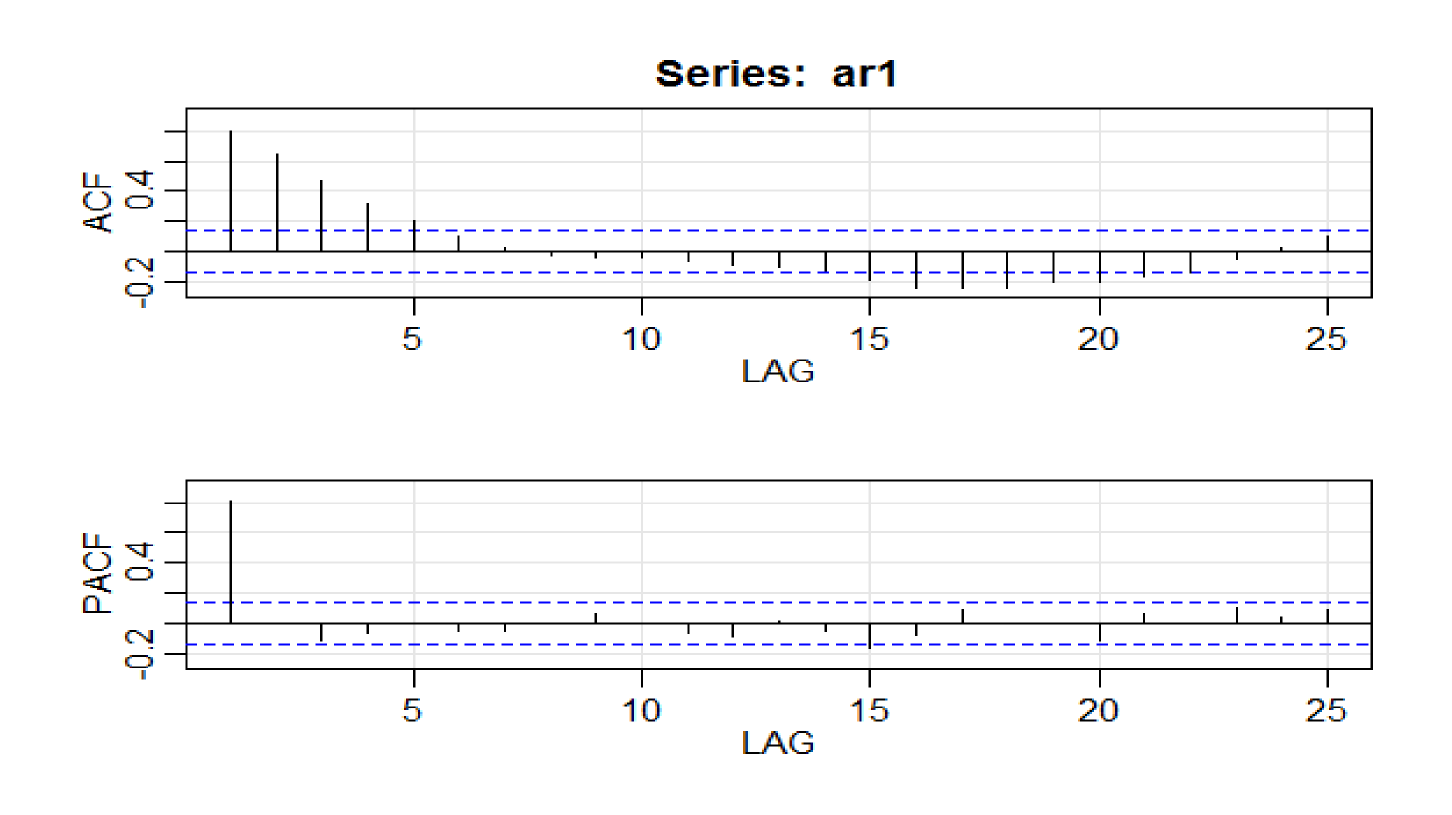
1. **Standardized Residuals**: Residuals should have a white noise
2. **ACF of Residuals**
3. **Normality Testing graph (Q-Q plot)**: Residuals should follow a Normal Distribution
4. **P-values of Ljung-Box Test**: All the points in the graph should be above the line

Thumb Rules for Identifying the Models

- We can guess the ARMA models with the help of ACF and PACF graphs
- Consider p = order of AR and q = order of MA in the table below
- Also note that
  - “Tails off after lag g” means Starting to decrease at g
  - “Cuts off after lag g” means Disappears or becomes very small at g

Graph	AR(p)	MA(q)	ARMA(p,q)
ACF	Tails off after lag p	Cuts off after lag q	Tails off after lag p
PACF	Cuts off after lag p	Tails off after lag q	Tails off after lag q

Example of AR(1)



- ACF is tailing off after lag 1 and PACF is cutting off after lag 1.
- Hence, the model can be AR(1) and we can try fitting ARMA(1,0,0) to the data.

Graph	AR(p)	MA(q)	ARMA(p,q)
ACF	Tails off after lag p	Cuts off after lag q	Tails off after lag p
PACF	Cuts off after lag p	Tails off after lag q	Tails off after lag q

Fitting of AR(1)

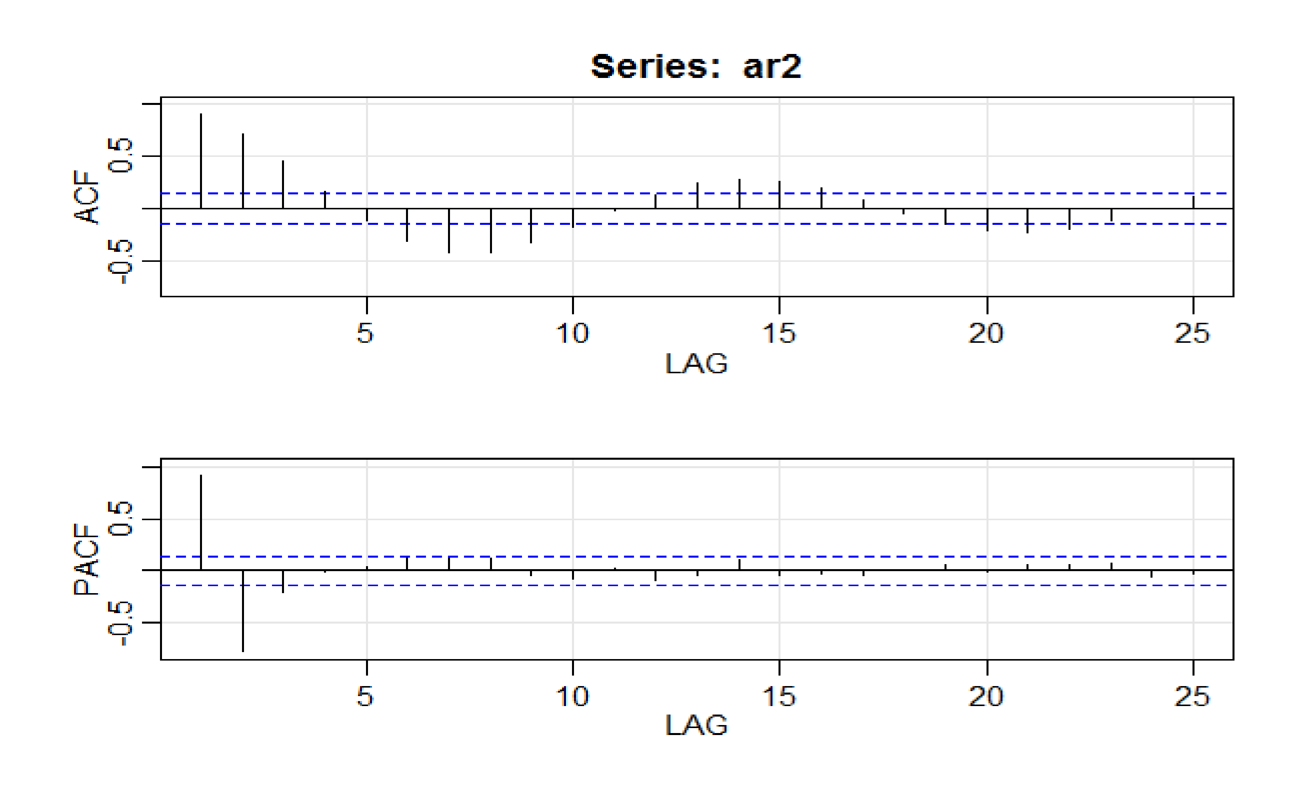
```
ar1.fit <- sarima(ar1,p=1,d=0,q=0)
ar1.fit$ttable
```

```
> ar1.fit$ttable
      Estimate     SE  t.value p.value
ar1     0.8038 0.0417  19.2705       0
xmean  40.5517 0.3781 107.2384       0
```

- The p-values in the ttable component indicate that the model is AR with order 1 with a shift of approximately 40.55. Hence, the model is (W_t is white noise at time t)

$$X_t = 40.5517 + 0.8038(X_{t-1} - 40.5517) + W_t$$

Example of AR(2)



- Though ACF is tailing off after lag 1 as well as after lag 2, PACF is cutting off after lag 2.
- Hence, the model can be AR(2) and we can try fitting ARMA(2,0,0) to the data.

Fitting of AR(2)

```
ar2.fit <- sarima(ar2,p=2,d=0,q=0)
ar2.fit$ttable
```

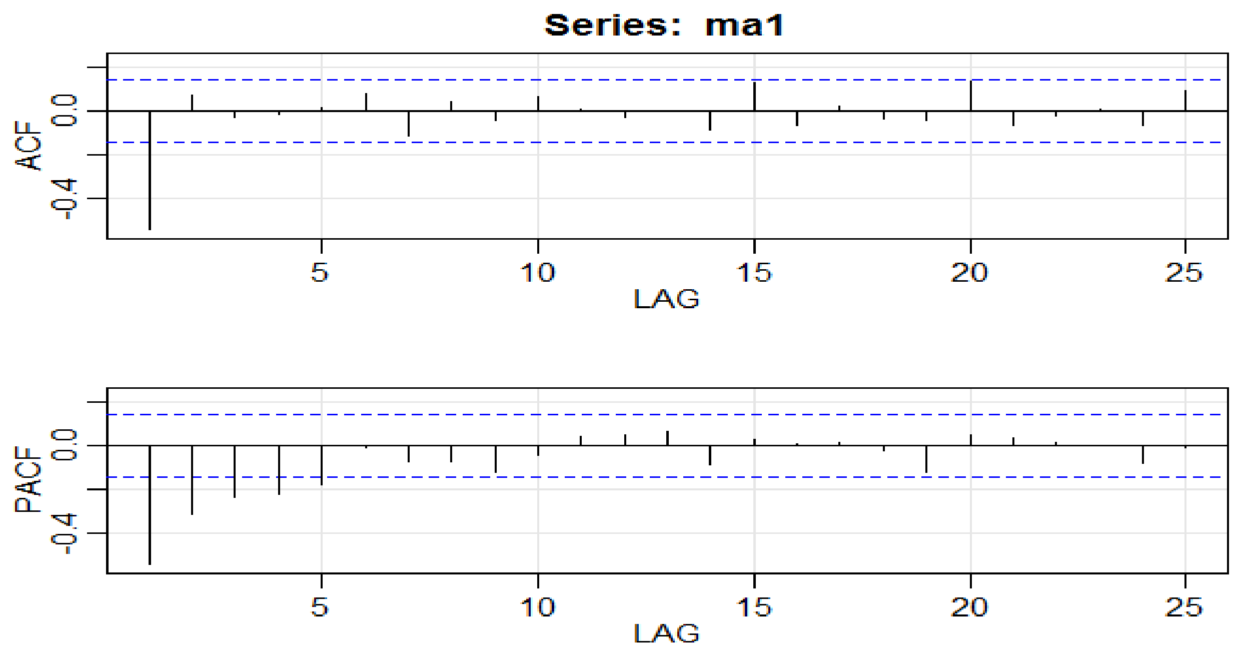
```
       Estimate     SE  t.value p.value
ar1      1.6649 0.0415  40.1525       0
ar2     -0.8093 0.0416 -19.4709       0
xmean   40.7873 0.4605  88.5765       0
```

- The p-values in the ttable component indicate that the model is AR with order 2 with a shift of approximately 40.7873. Hence, the model is (W_t is white noise at time t)

$$X_t = 40.7873 + 1.6649(X_{t-1} - 40.7873) - 0.8093(X_{t-2} - 40.7873) + W_t$$

Example of MA(1)

Graph	AR(p)	MA(q)	ARMA(p,q)
ACF	Tails off after lag p	Cuts off after lag q	Tails off after lag p
PACF	Cuts off after lag p	Tails off after lag q	Tails off after lag q



- ACF is cutting off after lag 1 and PACF is tailing off after lag 1.
- Hence, the model can be MA(1) and we can try fitting ARMA(0,0,1) to the data.

Fitting of MA(1)

```
ma1.fit <- sarima(ma1,p=0,d=0,q=1)
ma1.fit$ttable
```

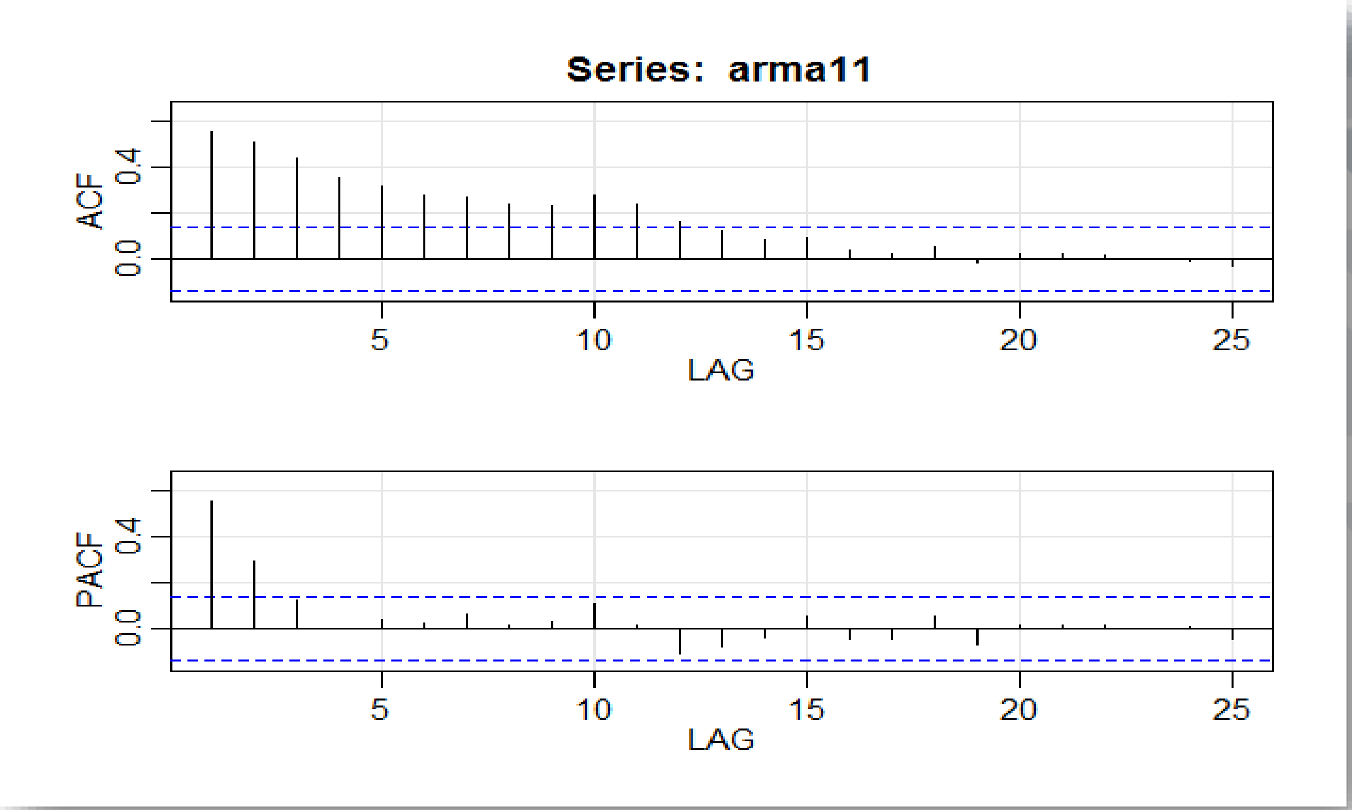
```
> ma1.fit$ttable
      Estimate     SE   t.value p.value
ma1    -0.8893 0.0350  -25.3893       0
xmean  40.0021 0.0076 5291.1493       0
```

- The p-values in the ttable component indicate that the model is MA with order 1 with a shift of approximately 40.7873. Hence, the model is (W_t is white noise at time t)

$$X_t = 40.0021 + W_t - 0.8893W_{t-1}$$

Example of ARMA(1,1)

Graph	AR(p)	MA(q)	ARMA(p,q)
ACF	Tails off after lag p	Cuts off after lag q	Tails off after lag p
PACF	Cuts off after lag p	Tails off after lag q	Tails off after lag q



- ACF is tailing off after lag 1 and PACF is tailing off after lag 1.
- Hence, the model can be ARMA(1,1) and we can try fitting ARMA(1,0,1) to the data.

Fitting of ARMA(1,1)

```r
arma11.fit <- sarima(arma11,p=1,d=0,q=1)
arma11.fit$ttable
```

```
      Estimate     SE  t.value p.value
ar1     0.8791 0.0493  17.8478  0.0000
ma1    -0.5058 0.0901  -5.6118  0.0000
xmean  -0.3402 0.2668  -1.2753  0.2037
```

- The p-values in the ttable component indicate that the model is AR with order 1 and MA with order 1. Hence, the model is (W_t is white noise at time t)

$$X_t = -0.3402 + 0.8791X_{t-1} + W_t - 0.5058W_{t-1}$$

Identifying the Models

- At first, while guessing, we start from smaller order and then gradually try for higher orders
- Model has to be chosen for which the AIC and BIC values are the least