Seasonal ARIMA models

- So far, we have considered the **ARMA** family of models, which rely on the assumption of stationarity.
- We now consider a more general family that allows the modeling of nonstationary time series through the application of differencing.
- The simplest example is the random walk example we discussed previously. Recall that, we defined the random walk $\{X_t\}$ as

$$X_t = X_{t-1} + w_t, \quad \text{where } \; w_t \sim \mathrm{WN}(0, \sigma^2)$$

$\{X_t\}$ is a nonstationary AR(1) process. However, $\{\nabla X_t\}$ with

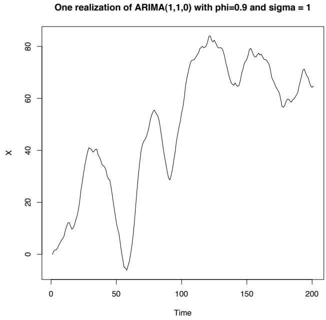
$$\nabla X_t = X_t - X_{t-1}$$

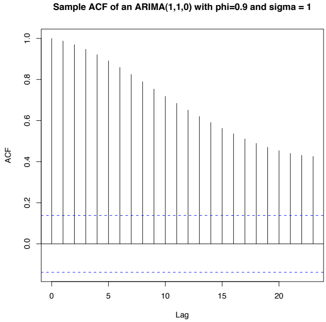
is a stationary process, being just the white noise w_t.

Simulated example

Simulate an **ARIMA**$(1,1,0)$ with $\phi = 0.9$ and $\sigma^2 = 1$.

```
X <- arima.sim(list(order = c(1,1,0), ar = 0.9), n = 200)
```





The ARIMA model

- The need for **ARIMA** arises from the fact that the series $\{X_t\}$ is nonstationary.
- Apply **differencing** operator $\nabla = 1 - B$ until the transformed series $\{Y_t = \nabla^d X_t\}$ exhibits stationarity.
- In the ARIMA equation

$$\phi^*(B)X_t \equiv \phi(B)(1-B)^d X_t = \theta(B)w_t, \quad w \sim \mathrm{WN}(0, \sigma^2) \quad (1)$$

  the integer $d \geq 0$ is the number of applications of differencing.
- Augmented Dickey-Fuller test

```
adfTest(X, lag=5)
  STATISTIC:Dickey-Fuller: 0.4971
  P VALUE:0.7737
```

Parameter Estimation

- Based on the original series X_t, we use $p = 1$, $d = 1$ and $q = 0$,

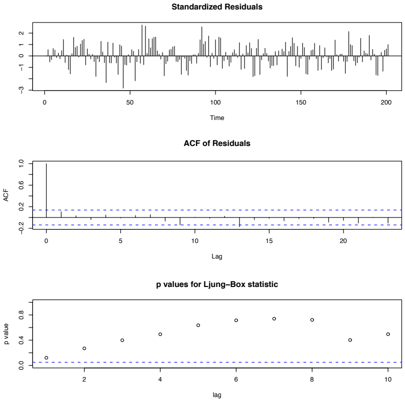
```
M1<-arima(X, order=c(1,1,0))
```

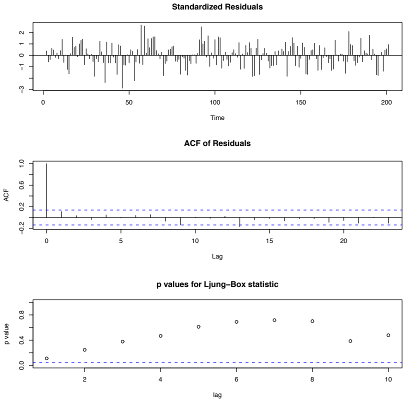
  We find $\hat{\phi} = 0.89$ with $\text{se}(\hat{\phi}) = 0.0318$ and $\hat{\sigma}^2 = 0.8924$.
- Based on the differenced $Y_t = \nabla X_t$, we use $p = 1$, $d = 0$ and $q = 0$,

```
M2<-arima(Y, order=c(1,0,0))
```

  We find $\hat{\phi} = 0.88$ with $\text{se}(\hat{\phi}) = 0.0322$ and $\hat{\sigma}^2 = 0.8906$.
- The diagnostics checks in both cases support the plausibility of the chosen model.

Diagnosis plots

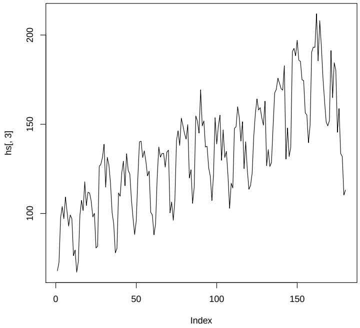


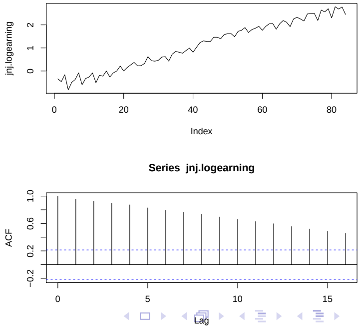


General procedure: Applying differencing until the resulting series has a sample ACF that decays rapidly, and the differenced data can be fitted by a low-order ARMA process.

Time series such as quarterly earnings and temperature index typically exhibits seasonal patterns.

```
hs=read.table("m-USTot.txt",header=T)
plot(hs[,3],type="l")
da=read.table("q-earn-jnj.txt")
jnj.logearning=log(da[,1])
plot(jnj.logearning,type="l")
```





Seasonal ARMA model

Formally, the **seasonal ARMA** model is defined by **ARMA**$(P, Q)_s$, with

$$\Phi(B^s)X_t = \Theta(B^s)w_t$$

where

$$\Phi(z) \quad = \quad 1 - \Phi_1 z - \Phi_2 z^2 - \cdots - \Phi_P z^P$$

and

$$\Theta(z) \quad = \quad 1 - \Theta_1 z - \Theta_2 z^2 - \cdots - \Theta_Q z^Q$$

are respectively, the **seasonal AR** operator and the **seasonal MA** operator, with period s.

Summary of ACF and PACF behaviors for seasonal ARMA

	AR$(P)_s$	**MA**$(Q)_s$	**ARMA**$(P, Q)_s$
ACF	Tails off at lags ks	Cuts off after lags Qs	Tails off at lags ks
PACF	Cuts off after lags Ps	Tails off at lags ks	Tails off at lags ks

- The values of ACF and PACF are zero at non-seasonal lags $\tau \neq ks$.
- Here, s is the length of the period, and $k = 1, 2, \cdots$.

Mixed Seasonal ARMA Models

- The seasonal and non-seasonal ARMA models can be combined. The resulting model has equation

$$\Phi(B^s)\phi(B)X_t = \Theta(B^s)\theta(B)w_t$$

- This resulting model is called the **mixed seasonal ARMA**, and is denoted by

$$\textbf{ARMA}(p, q) \times (P, Q)_s$$

- The behavior of a **mixed seasonal ARMA** model is a combination of the the behaviors of its seasonal and nonseasonal constituents.

- In practice, the value of P and Q are typically less than 3.
- Johnson and Johnson quarterly earning example
- Let's take a look at the ACF of J&J log earning series.

```
par(mfcol=c(2,1))
acf(jnj.logearning,lag.max=16)
jnj.return=diff(jnj.logearning)
acf(jnj.return)
```

JJ earnings

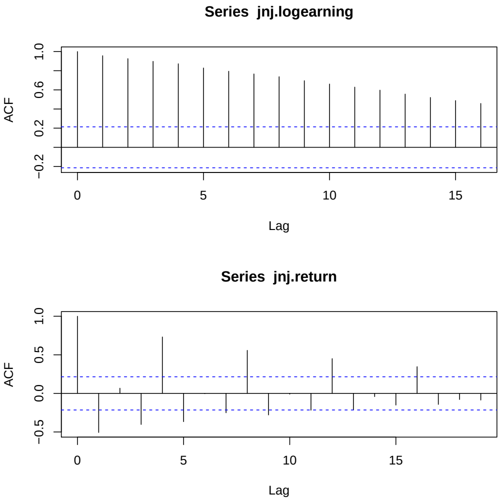


Figure: ACF of JJ log earning and returns.

Seasonal ARIMA model for JJ data

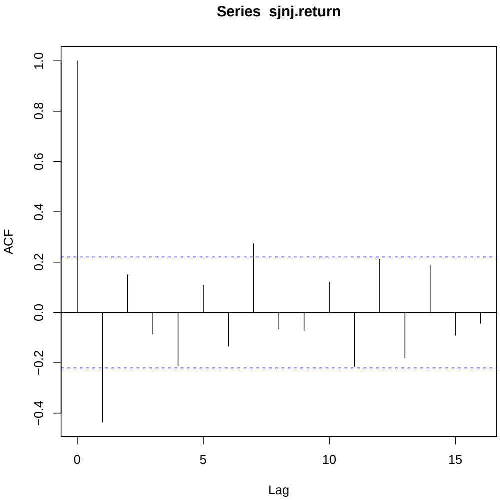
- Clearly, the original series seems to be nonstationary. So does the regular differenced series.
- The airline model (for quarterly series):

$$(1-B)(1-B^4)w_t = (1-\theta_1 B)(1-\theta_4 B^4)w_t.$$

  After regular and seasonal differencing, this is essentially a is a multiplicative MA model.
- The left hand side can be written as $Y_t = (w_t - w_{t-1}) - (w_{t-4} - w_{t-5})$. We call it regular and seasonal differenced series.
- Now perform seasonal differencing to the jnj.reutrn series.

Seasonal and regular differenced JJ data

```
sjnj.return=diff(jnj.return,4)    ## yearly = 4 quarters
acf(sjnj.return,lag.max=16)
```

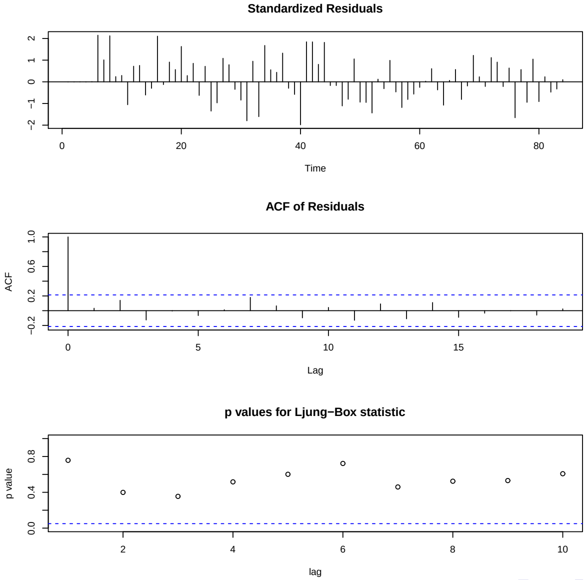


Airline model

Apply the airline model to jnj logearning.

```
##define the airline model:

m1=arima(jnj.logearning,order=c(0,1,1)
,seasonal=list(order=c(0,1,1),period=4))

# This model has aic = -150.75
# If we get rid of seasonal MA term,  aic = -145.51

tsdiag(m1) # Model checking

f1=predict(m1,8) # prediction for 2 years
```

Diagnosis plots



Out of sample prediction