ARIMA model is the same as Box-Jenkins model

ARIMA is perfect for working with univariate data.However, since arima is flexible and thus very general model sometimes exponential smoothing might be a better fit, especially seasonal datasets are easier to  communicate with exponential smoothing or seasonal decomposition.

AR-Autoregressive part:p

I- Integration, degree of differencing:d

MA- moving average part:q

These 3 parameters are integers that describe the grade or the order of the three parts of the AR-I-MA model. Arima model requires AR and MA parts to be stationary time series, the model will make it stationary by differencing done by the model function. how often or the order of differencing is described by the d part of the ARIMA(p,d,q)

you can manually do the differencing with diff() or you can let the model do it for you.

How to calculate the three parameters?

1. arima() -estimating the parameters manually by using the ACF and PACF plots---R Base

2. auto.arima()- R calculates the parameters automatically and chooses a suitable model- library(forecast)

How to read an ARIMA model?

The whole model is based on:

-summation of lags = autoregressive part.

-summation of forecasting errors = moving average part

-coefficient: determines the importance of a specific lag

For example:

if a model only has autoregressive single lag component:

we denote it as AR(1) or ARIMA(1,0,0): first order (lag) of AR

AR(2) or ARIMA(2,0,0) second order of AR

MA(1) or ARIMA(0,0,1): first order of MA

How to Calculate a model? – parameters p and q

1.**P**: AR(1) or ARIMA(1,0,0) – yt = δ + φ1yt-1 + 𝜀t

This reads: the observed value(Yt) at time point t consist of

* The constant δ
* The value of the previous point(Yt-1) multiplied by a coefficient φ
* The error term of time point t(𝜀t)

2. **Q**:ARMA(1,1) or ARIMA(1,0,1) - yt = δ + φ1yt-1 + Θet-1 + 𝜀t

- Extra step: forecast error term for the first lag(**e**t-1) multiplied by coefficient Θ

- Forecast error at t: the difference between the actual and the forecasted value at time time

3. How to calculate parameter d?

-Example: ARIMA(0,1,0) Random walk: The mean is not constant(not stationary), which is required for a forecast.

Chart

Description automatically generated

-The solution here is differencing: The present observation or expected value at time t minus the previous observation equals a constant plus an error term at time point t. yt - yt-1 = δ + 𝜀t

Often time backshift operator: B or L represents a differenced time series

Byt =  yt-1