ARIMA 모델

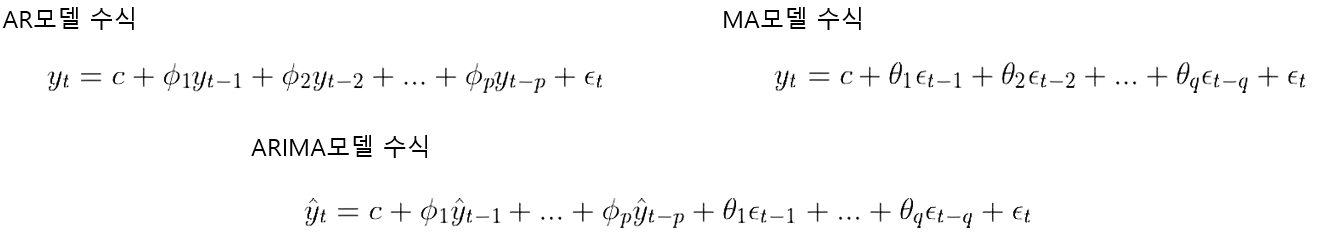
# 개요

* ARIMA는 Auto Regressive Intergrated Moving Average로 시계열 데이터의 패턴을 파악하여 통계적 기법으로 미래값을 예측하는 모델이다. ARIMA모델 이전 예측모델인 AR과 MA를 합친 후 차분하는 과정을 추가한 모델로 이전의 모델들과는 달리 비정상성을 가진 데이터도 차분을 통해 정상성을 가지게 한 후 예측이 가능하게 하였다.

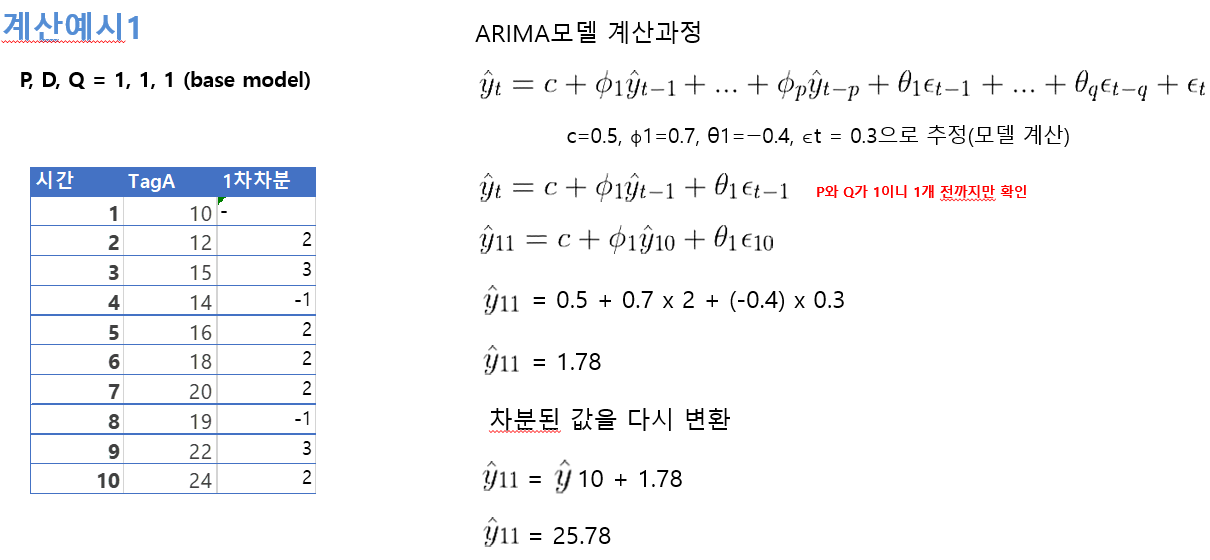
# 구성요소 및 파라미터

* 구성요소
  + AR(Auto Regressive) : 과거 값들이 현재 값에 미치는 영향
  + I(Intergrated) : 시계열 데이터가 정상성을 가질 수 있게 차분을 적용
  + MA(Moving Average) : 과거 오차들이 현재 값에 미치는 영향
* 파라미터
  + P : 자기회귀(Auto Regressive)되는 항의 수를 의미   
     이전의 관측 값들이 현재의 값에 얼마나 영향을 미치는지 나타냄
  + D : 차분 차수를 의미  
     시계열의 비정상성을 제거하고 데이터의 평균이나 분산이 시간에 따라 일정하도록 함  
     시계열 데이터의 추세(Trend)를 제거하기 위해 사용됨
  + Q : 이동평균(Moving Average)되는 항의 수를 나타냄  
     예측 값이 예측 오차에 따라 조정되는 정도를 의미하고 시계열데이터의 잡음을   
     모델링하기 위해서는 Q값을 설정하여 오차정도를 보정할 수 있

# 수식

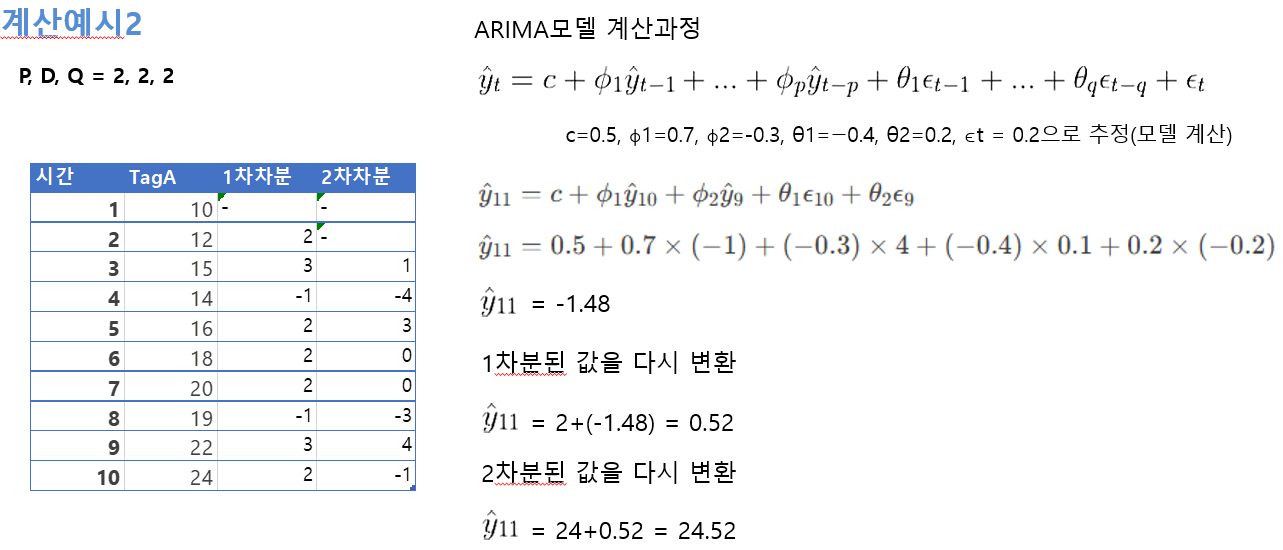
* 
* ARIMA의 수식은 기존 예측모델인 AR모델과 MA모델의 수식을 합친 모양으로   
  c, ϕ, θ는 모델에 넣는 학습데이터에 의해 정해지게 되며 p, q가 높아질수록 모델의 복잡성은 증가하고  
  d차분의 경우 비정상성을 없애기 위한 방법으로 최대 2차 차분까지만 하는것이 바람직하다.

# 계산과정 예시



* Step size를 늘려 더 미래의 예측을 진행할때는 기존에 예측했던 11번째 데이터까지

사용하여 파라미터를 추정하고 예측하는 재귀적 예측이 수행되게됨



* P와 Q를 통해 얼마이전의 값까지 사용할 것인지 정하고 2차차분을 했을 때는 다시 반환하는 과정도 2번 수행되게 됨