1. AR(p) 모형
2. MA(q) 모형
3. ARMA(p,q) 모형

## (1) 모형

AR(1) $y_t = \emptyset_1 y_{t-1} + e_t$ , $e_t \sim iidN(0, \sigma_e^2)$

AR(p) $y_t = \mu + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} + e_t$ , $e_t \sim iidN(0, \sigma_e^2)$

## (2) AR(1)

① ACF

– 반복적인 대입을 통해 $y_t$를 교란항의 함수로만 표현해 보면($y_0 = 0$)

$$
\begin{aligned}
y_t &= \phi y_{t-1} + e_t, \quad e_t \sim (0, \sigma_e^2) \\
&= \phi^2 y_{t-2} + \phi e_{t-1} + e_t \\
&\quad \vdots \\
&= \sum_{j=0}^{\infty} \phi^j e_{t-j}
\end{aligned}
$$

- $y_t$의 평균, 분산 및 공분산을 구해 보면 다음과 같음

$$E(y_t) = 0$$

$$Var(y_t) = \sigma_e^2 + \phi^2\sigma_e^2 + \phi^4\sigma_e^2 + .......$$

$$= \frac{\sigma_e^2}{1-\phi^2}$$

$$\equiv \gamma_0$$

$$Cov(y_t y_{t-1}) = E(y_t y_{t-1}) = E[(e_t + \phi e_{t-1} + ...)(e_{t-1} + \phi e_{t-2} + ...)]$$

$$= \phi\sigma_e^2 + \phi^3\sigma_e^2 + ...$$

$$= \frac{\phi}{1-\phi^2}\sigma_e^2 \equiv \gamma_1$$

$$Cov(y_t y_{t-k}) = \frac{\phi^k}{1-\phi^2}\sigma_e^2 \equiv \gamma_k$$

– 따라서 다음과 같이 $\boldsymbol{\rho_0}$, $\boldsymbol{\rho_1}, \cdots, \boldsymbol{\rho_k}$를 구할 수 있음

$$\therefore \quad \rho_0 \equiv \frac{\gamma_0}{\gamma_0} = 1$$

$$\rho_1 \equiv \frac{\gamma_1}{\gamma_0} = \phi$$

.

$$\rho_k \equiv \frac{\gamma_k}{\gamma_0} = \phi^k$$

– 이를 그래프로 나타낸 것을 상관도(correlogram)라 하며 이를 이용하여 시계열의 안정성여부를 판단할 수 있음

– 시계열이 안정적이면(ø→0), ACF는 급격하게 0을 향해 감소하고

시계열이 불안정적이면(ø→1), ACF는 0을 향해 천천히 감소함

② PACF

- k 이외의 모든 시차를 갖는 관측치($y_{t-1}$, $y_{t-2}, \cdots, y_{t-k+1}$)의 영향력을 배제한 가운데 특정의 두 관측치, $y_t$ 및 $y_{t-k}$가 얼마나 관련이 있는 지 나타내는 척도로 회귀계수가 편자기상관함수가 됨
- 즉, $\phi_{kk} = corr(y_t, y_{t-k} \mid y_{t-1}, y_{t-2}, \cdots, y_{t-k+1})$
- 시차 k의 편자기상관계수는 다음 식에서 k번째 회귀계수 $\phi_{kk}$를 의미함

$$y_t = \phi_{k1} y_{t-1} + \phi_{k2} y_{t-2} + \ldots + \phi_{kk} y_{t-k} + e_t$$

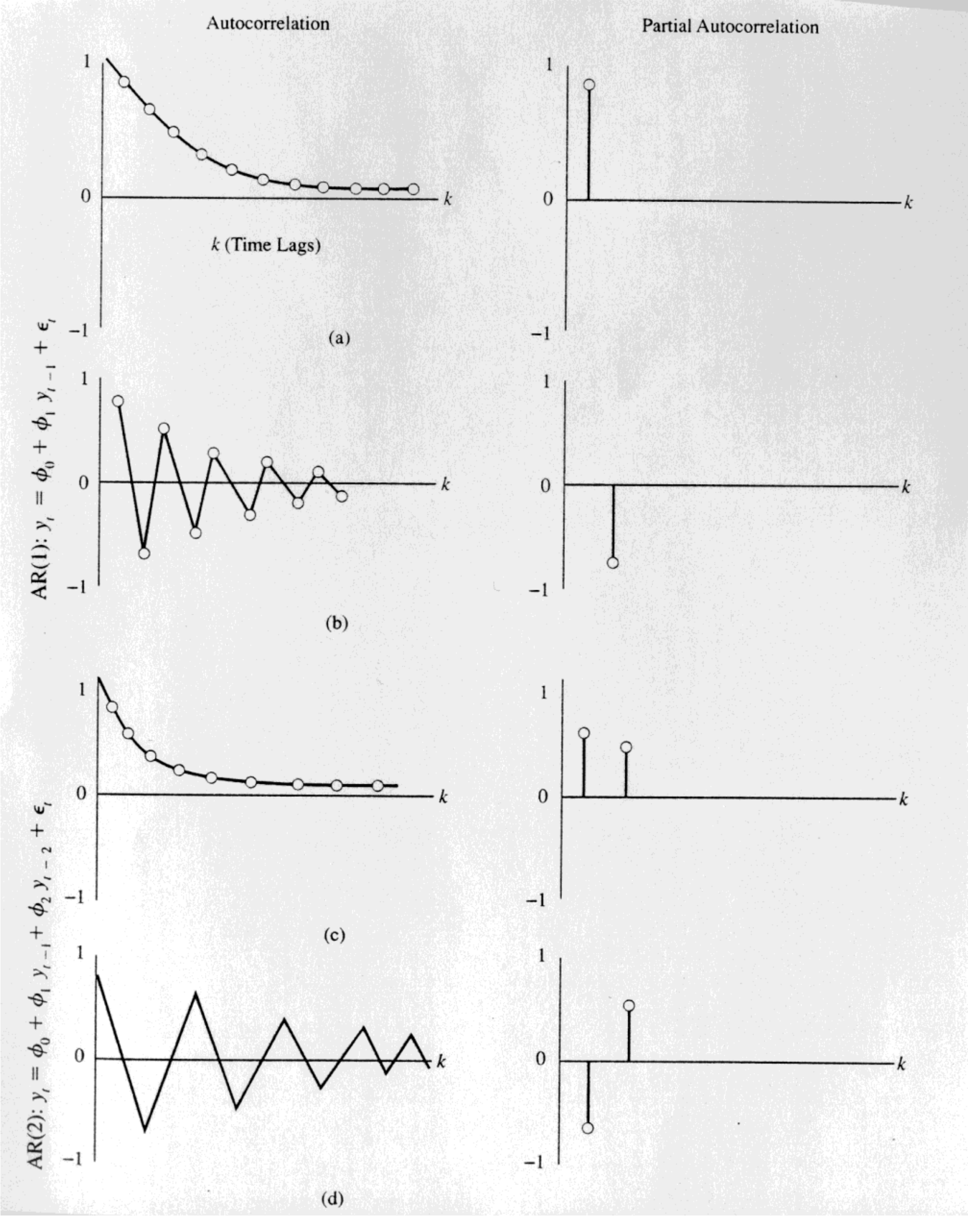
즉, AR(1): $\phi_{11} = \phi_1$

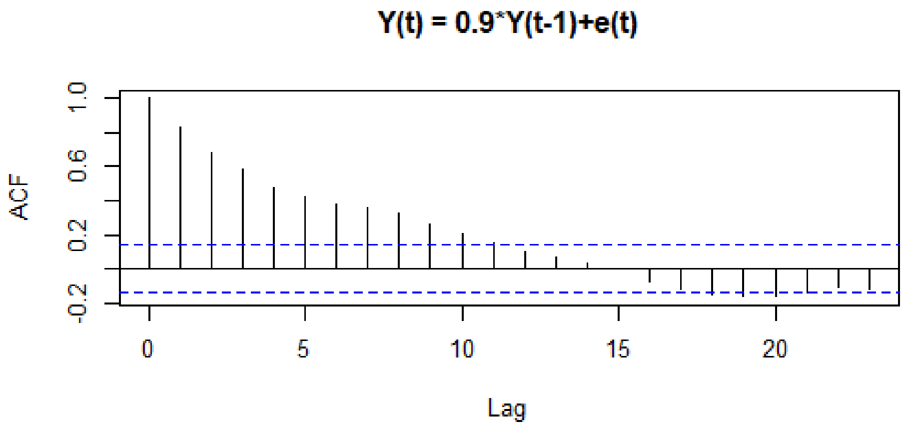
AR(2): $\phi_{22} = \phi_2$

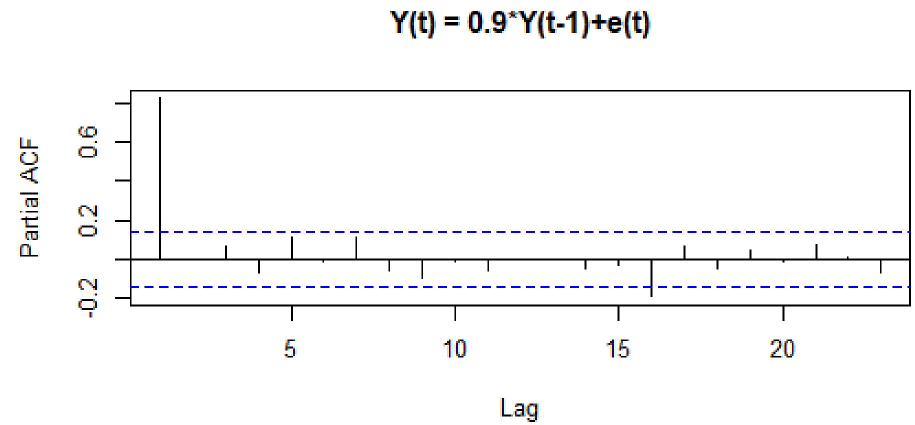
③ 식별

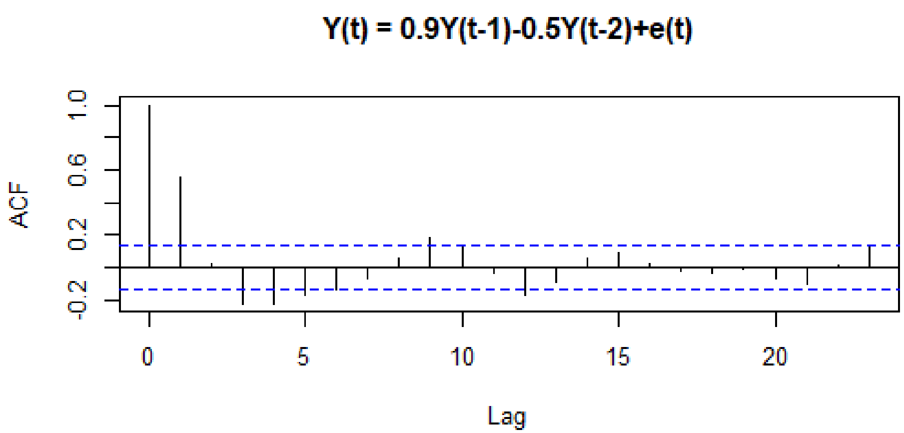
- ACF가 점진적으로 감소하면 불안정시계열이므로 원계열을 차분하여 안정시계열로 만들어 줌

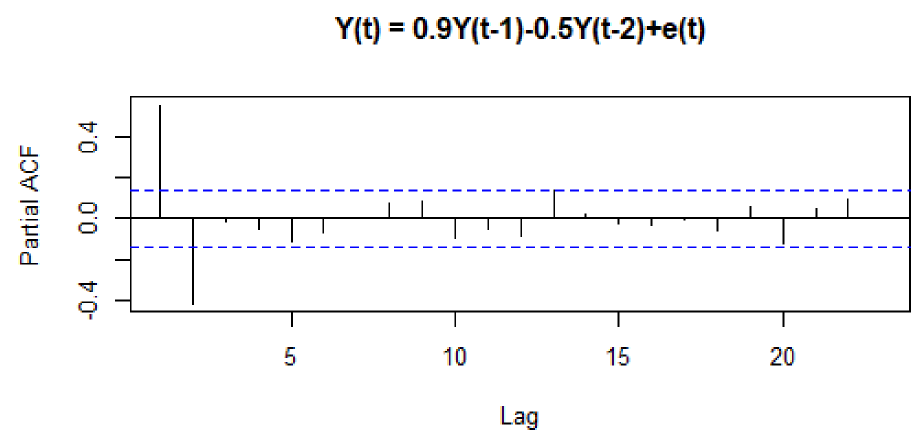
- ACF가 0을 향해 감소하고 PACF는 1-2개 정도만 통계적으로 유의하면 AR모형




Y(t) = 0.9*Y(t-1)+e(t)
ACF
-0.2
0.2
0.6
1.0
0
5
10
15
20
Lag


Y(t) = 0.9*Y(t-1)+e(t)
Partial ACF
-0.2
0.2
0.6
5
10
15
20
Lag


Y(t) = 0.9Y(t-1)-0.5Y(t-2)+e(t)
ACF
-0.2
0.2
0.6
1.0
0
5
10
15
20
Lag


Y(t) = 0.9Y(t-1)-0.5Y(t-2)+e(t)
Partial ACF
-0.4
0.0
0.4
5
10
15
20
Lag

④ 최적 모형의 선정(AIC 기준)

- ACF 및 PACF를 이용하여 모형을 식별할 수도 있지만 최적 모델을 선정하는 기준으로 적합성을 측정해 주는 AIC(Akaike Information Criterion) 또는 SBC(Schwarz-Bayesian Criterion) 등이 있음
- AIC 및 SBC는 다음과 같이 계산되는데 작은 값일수록 모형의 적합도가 높음을 나타냄

$$AIC = 2*k - 2*\ln(L^*)$$

$$SBC = k*\ln(n) - 2*\ln(L^*)$$

단, k는 추정계수의 개수, n은 관측치 수, $\ln(L^*)$은 최적화된 log likelihood의 값

⑤ 추정

- 일반적으로 AR 모형은 최소자승법으로 추정하고 MA 모형은 비선형최소자승법으로 추정
- 두 모형 모두 최우법으로 추정하기도 함

```
> ar1.sim<-arima.sim(model=list(ar=c(0.7)), n=100)

> ar1<-arima(ar1.sim, order=c(1,0,0), include.mean=T)

> ar1

Call:
arima(x = ar1.sim, order = c(1, 0, 0), include.mean = T)

Coefficients:
         ar1  intercept
      0.7127     0.0226
s.e.  0.0685     0.2947

sigma^2 estimated as 0.7526:  log likelihood = -128.04,  aic = 262.08
```

(주의사항)

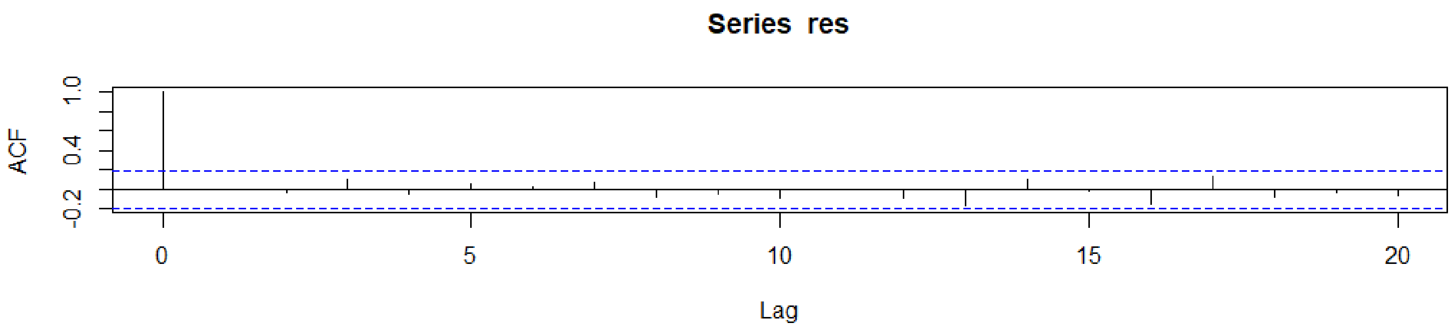
-$y_t = \mu + \phi y_{t-1} + e_t$는 $(1-\phi L)y_t = \mu + e_t$이므로 $y_t = \frac{\mu}{1-\phi} + (1-\phi L)^{-1} e_t$로 나타낼 수 있고

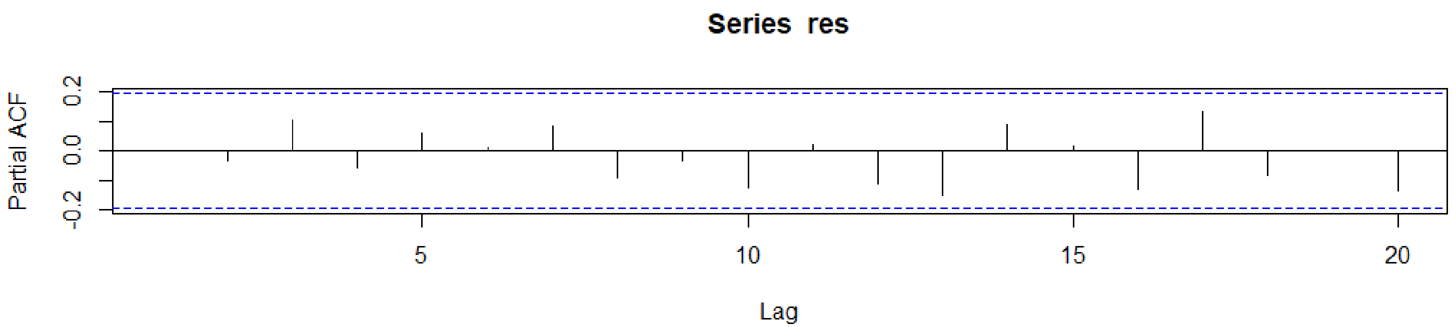
AR(1)에서 추정된 상수항 $\hat{c} = \frac{\hat{\mu}}{1-\hat{\phi}}$이므로 $\hat{\mu} = \hat{c}(1-\hat{\phi})$이 됨

-따라서, $\hat{\phi} = 0.7127$이고, $\hat{c} = 0.0226$이므로 $\hat{\mu} = 0.02269(1-0.7127) = 0.0065$임.

## ⑥ 모형의 검토

- 잔차분석 : 모형설정이 제대로 되었으면 잔차항은 백색잡음확률과정을 따를 것임




- 잔차의 독립성 검정 : Ljung-Box 검정

$$Q = T(T+2)\sum_{k=1}^{K}\frac{1}{T-k}\gamma_k^2 \sim \chi^2_{K-p-q}$$

단, $K = Min(\frac{T}{2}, T^{1/3})$

```
> Box.test(res, lag=1, type=c("Ljung-Box"))

        Box-Ljung test

data:  res
X-squared = 1.5786e-05, df = 1, p-value = 0.9968
```

$H_0 : \rho_1 = \rho_2 = ... = \rho_k = 0$ (단, $\rho_k$ (k=1,2,3...K)는 잔차의 acf임)

## ⑦ 예측

- 추정된 AR(p)모형을 이용하면 미래의 값에 대한 예측치를 도출할 수 있는데 AR(1)의 경우를 예를 들어 n시점까지의 정보를 바탕으로 k기 이후 즉, (n+k)기에 대한 예측치를 구하는 방법은 다음과 같음

$$y_t = \mu + \phi y_{t-1} + e_t$$

- 위 식에서 (n+1)기(k=1)의 $y$ 즉, $y_{n+1}$은 다음과 같이 나타낼 수 있음

$$y_{n+1} = \mu + \phi y_n + e_{n+1}$$

- 따라서 n시점까지의 정보를 이용하여 (n+1)기에 대한 예측치를 구하면 다음과 같음.

$$\hat{y}_{n+1} = \hat{\mu} + \hat{\phi} y_n \quad \text{또는} \quad E[y_{n+1} \mid I_t] = \mu + \phi y_n \leftarrow E[e_{n+1} \mid I_t = 0]$$

- 따라서, n시점까지의 정보를 이용하여 (n+1)기 예측치에 대한 예측오차(실제치-예측치)와 그 분산은 다음과 같음

- 예측오차 :

$$\begin{aligned} & y_{n+1} - E[y_{n+1} \mid I_t] \\ &= \mu + \phi y_n + e_{n+1} - (\mu + \phi y_n) \\ &= e_{n+1} \end{aligned}$$

- 예측오차 분산 :

$$\begin{aligned} & Var(y_{n+1} - E[y_{n+1} \mid I_t]) \\ &= Var(e_{n+1}) \\ &= \sigma_e^2 \end{aligned}$$

- 유사한 방법으로 n시점에서의 (n+2)기에 대한 예측치는 다음과 같음

$$\hat{y}_{n+2} = \hat{\mu} + \hat{\phi} y_{n+1}$$

$$= \hat{\mu} + \hat{\phi}(\hat{\mu} + \hat{\phi} y_n)$$

$$= (\hat{\phi}+1)\hat{\mu} + \widehat{\phi^2} y_n$$

또는 $E[y_{n+2} \mid I_t] = \mu + \phi E[y_{n+1} \mid I_t] \leftarrow E[e_{n+2} \mid I_t = 0]$

- 따라서, n시점까지의 정보를 이용하여 (n+2)기 예측치에 대한 예측오차(실제치-예측치)와 그 분산은 다음과 같음

- 예측오차 : $y_{n+2} - E[y_{n+2} \mid I_t]$

$$= \mu + \phi y_{n+1} + e_{n+2} - \mu - \phi E[y_{n+1} \mid I_t]$$
$$= \phi (y_{n+1} - E[y_{n+1} \mid I_t]) + e_{n+2}$$
$$= \phi e_{n+1} + e_{n+2}$$

- 예측오차 분산 : $Var(y_{n+2} - E[y_{n+2} \mid I_t])$

$$= Var(\phi e_{n+1} + e_{n+2})$$
$$= (\phi^2 + 1)\sigma_e^2$$

○ 일반적으로 n시점에서 h기 이후 즉, (n+h)기에 대한 예측치는 다음과 같음

$$\hat{y}_{n+h} = (\hat{\phi}^{h-1} + \hat{\phi}^{h-2} + \dots + 1)\hat{\mu} + \hat{\phi}^h y_n$$

```
 [1]  0.17961519 -0.34706078 -1.31076625 -1.13551129 -1.82086235 -2.00349488 -2.02748568 -3.10593329
 [9] -1.33636626 -0.78208326 -1.68559522  0.07389827  0.47819301  0.03966362  0.92289020  1.52415663
[17]  1.88849072  2.01058376  1.96132628  1.31101669  0.61174902  0.04775331 -0.66127966 -0.67081304
[25] -1.73496548  0.95448013  1.87609809  0.19016008 -0.26977278 -0.65549630  0.32111771  0.14141333
[33]  0.35230784  0.21806874  0.10977766  1.44544664  0.78604167  2.06669977 -0.10206297  0.51316967
[41]  0.48307302  0.55409268  0.76750436  0.03492960 -0.30875667 -1.23470505 -1.93608476 -1.05173069
[49] -0.28800171 -0.14859697  0.81824959  2.62285940  1.34497041 -1.36768959  0.04835581 -0.67535169
[57] -1.16075480  0.21304301 -0.13564290 -1.31566774 -0.73966394 -0.65665612 -0.45389510  0.06755383
[65] -0.32337235  0.41801590  0.07212457  0.38226916  1.36442743  1.39028069  0.64726490  1.60189305
[73]  2.11482899  2.02877725  1.65887581  0.53330699  1.73396734  0.61351755  2.61679528  3.36436732
[81]  2.11935677  0.45712884 -0.39041638 -0.01640776 -0.25817731 -0.52826671 -1.32140527 -0.97001141
[89] -1.46391246 -2.69268066 -2.26510298 -0.66657548 -1.04194980 -0.12140054 -1.70286308 -1.24756612
[97] -0.35388908  0.05343100  0.14307790 -0.54055148
```

(점 예측치)

-n시점에서의 (n+1)기에 대한 예측치

$$\hat{y}_{t+1} = \hat{\mu} + \hat{\phi} y_t = 0.0065 + (0.7127)(-0.5405) = -0.3787$$

-n시점에서의 (n+2)기에 대한 예측치

$$\hat{y}_{t+2} = \hat{\mu} + \hat{\phi} y_{t+1} = \hat{\mu} + \hat{\phi}(\hat{\mu} + \hat{\phi} y_n) = 0.0065 + (0.7127)(-0.3787) = -0.2634$$

```
> predict(ar1,n.ahead=4)
$pred
Time Series:
Start = 101
End = 104
Frequency = 1
[1] -0.3787392 -0.2634227 -0.1812416 -0.1226748

$se
Time Series:
Start = 101
End = 104
Frequency = 1
[1] 0.867547 1.065311 1.152832 1.194831
```

(구간예측)

-n시점에서의 (n+1)기에 대한 95% 구간예측

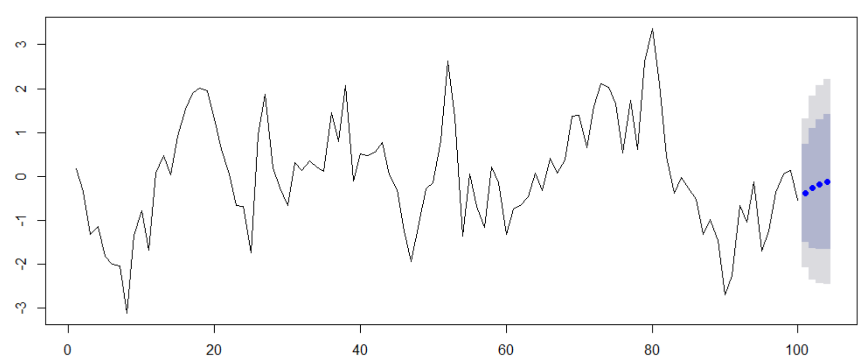
$$-0.3787 \pm (1.96)\sqrt{0.7526} = [-2.079,\ 1.3216]$$

-n시점에서의 (n+2)기에 대한 95% 구간예측

$$-0.2634 \pm (1.96)\sqrt{(1+0.7127^2)(0.7526)} = [-2.3513,\ 1.8245]$$

```
> (far1<-forecast(ar1,h=4))
    Point Forecast     Lo 80     Hi 80     Lo 95    Hi 95
101     -0.3787392 -1.490545 0.7330671 -2.079100 1.321622
102     -0.2634227 -1.628673 1.1018278 -2.351393 1.824548
103     -0.1812416 -1.658656 1.2961726 -2.440752 2.078268
104     -0.1226748 -1.653912 1.4085622 -2.464500 2.219150
```

Forecasts from ARIMA(1,0,0) with non-zero mean



## (1) 모형

MA(1) $y_t = e_t - \theta e_{t-1}$ , $e_t \sim iidN(0, \sigma_e^2)$

MA(q) $y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$ , $e_t \sim iidN(0, \sigma_e^2)$

## (2) MA(1)

### ① ACF

- $y_t$의 평균, 분산, 공분산 및 ACF를 구해 보면 다음과 같음

$$E(y_t) = 0$$

$$Var(y_t) = \sigma_e^2 + \theta^2 \sigma_e^2 = (1+\theta^2)\sigma_e^2 \equiv \gamma_0$$

$$Cov(y_t y_{t-1}) = E(y_t y_{t-1}) = E[(e_t - \theta e_{t-1})(e_{t-1} - \theta e_{t-2})] = -\theta \sigma_e^2 \equiv \gamma_1$$

$$Cov(y_t y_{t-2}) = E(y_t y_{t-2}) = E[(e_t - \theta e_{t-1})(e_{t-2} - \theta e_{t-3})] = 0 \equiv \gamma_2$$

$$\therefore \quad \rho_1 \equiv \frac{\gamma_1}{\gamma_0} = \frac{-\theta}{1+\theta^2}$$

$$\rho_2 \equiv \frac{\gamma_2}{\gamma_0} = 0$$

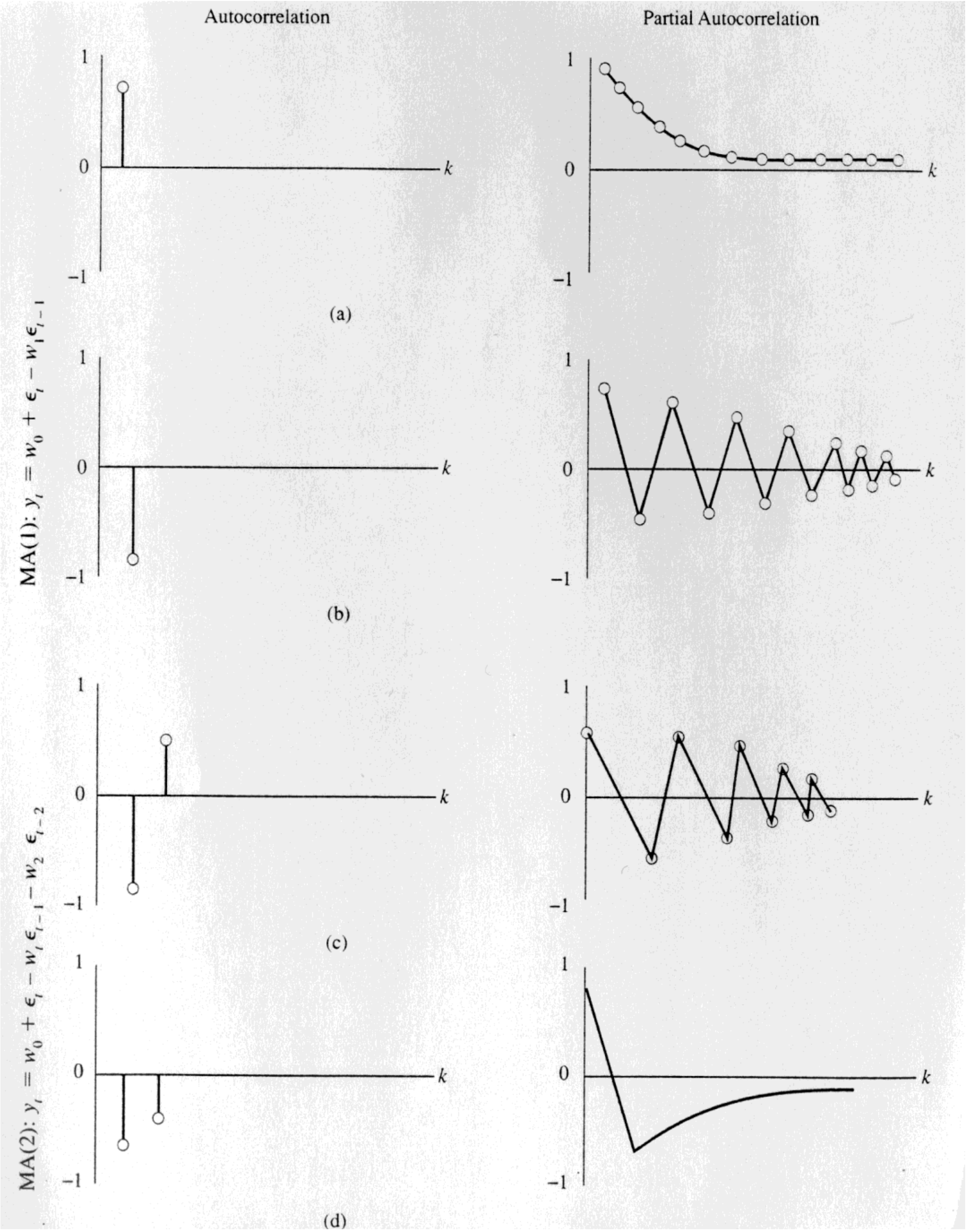
.

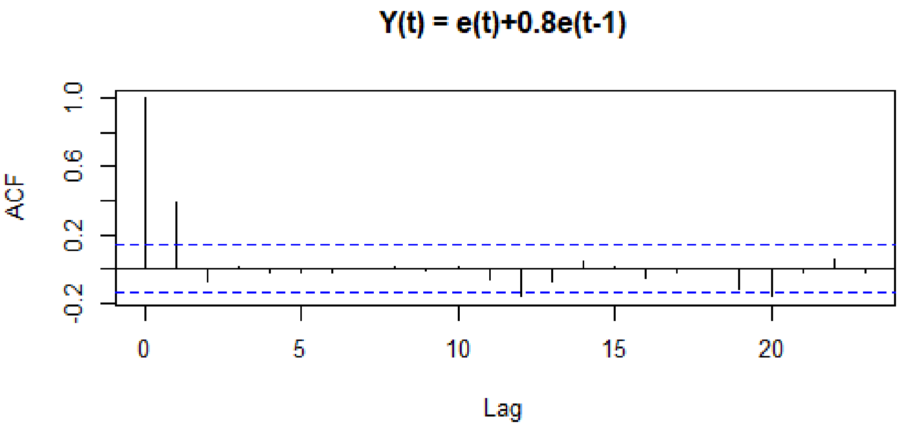
.

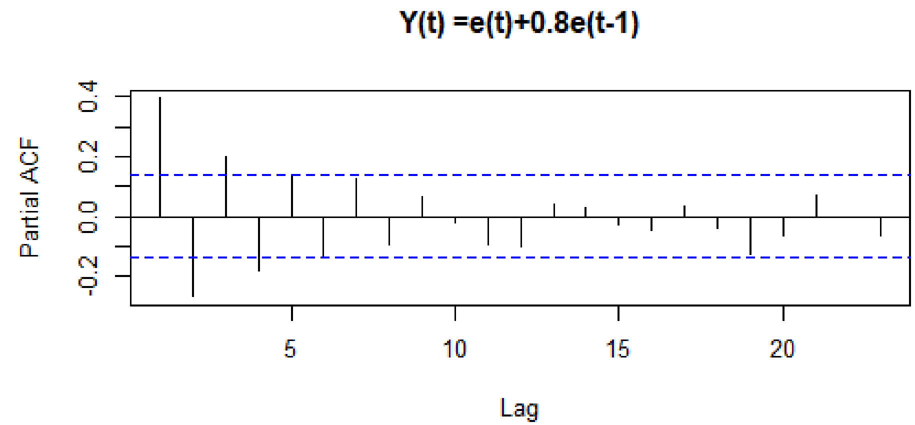
$$\rho_k = 0$$

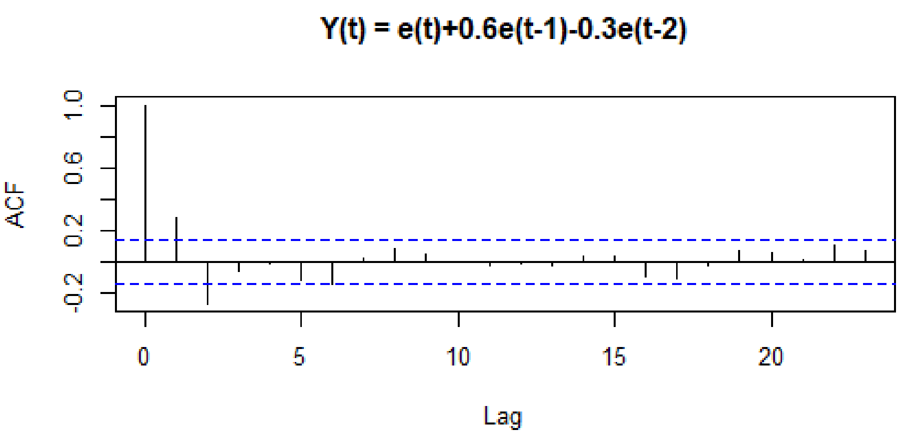
② 식별

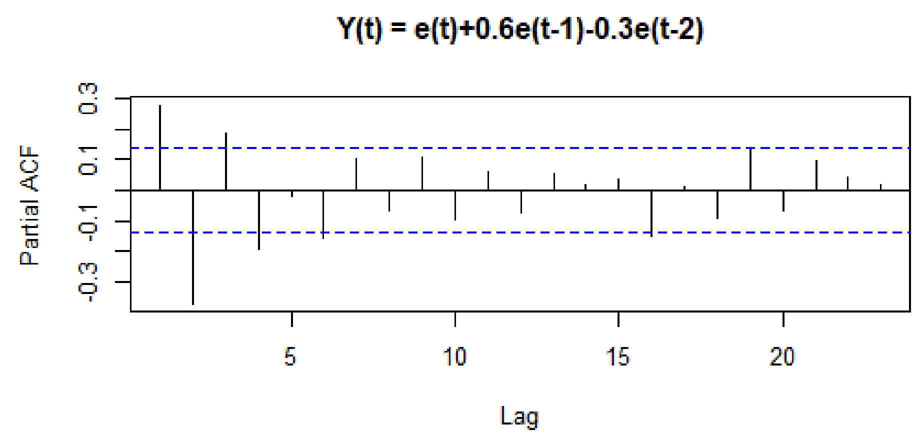
- PACF는 0을 향해 감소하고 ACF는 1 – 2개 정도만 통계적으로 유의하면 MA모형




Y(t) = e(t)+0.8e(t-1)
ACF
Lag


Y(t) =e(t)+0.8e(t-1)
Partial ACF
Lag


Y(t) = e(t)+0.6e(t-1)-0.3e(t-2)
ACF
Lag

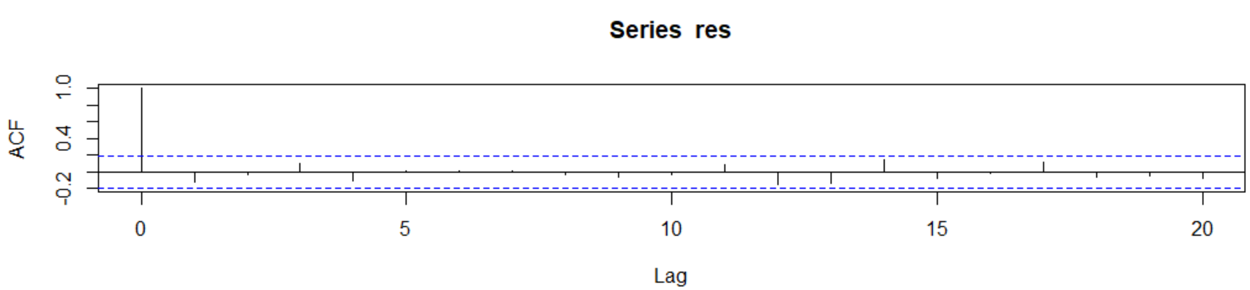

Y(t) = e(t)+0.6e(t-1)-0.3e(t-2)
Partial ACF
Lag

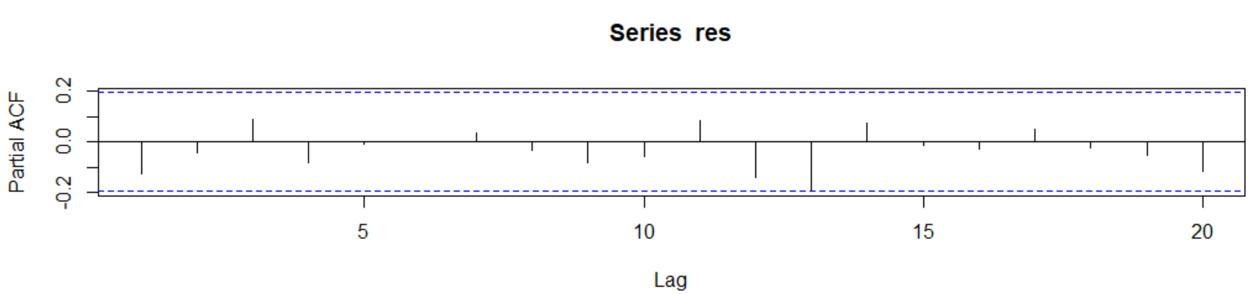
## ③ 추정

```
> ma1<-arima(ma1.sim, order=c(0,0,1), include.mean=F)

> ma1

Call:
arima(x = ma1.sim, order = c(0, 0, 1), include.mean = F)

Coefficients:
         ma1
      0.7399
s.e.  0.0858

sigma^2 estimated as 0.8194:  log likelihood = -132.33,  aic = 268.67
```

## ④ 모형의 검토

– 잔차분석 : 모형설정이 제대로 되었으면 잔차항은 백색잡음확률과정을 따를 것임



⑤ 예측

– 추정된 MA(q)모형을 이용하면 미래의 값에 대한 예측치를 도출할 수 있는데 MA(1)의 경우를 예를 들어 n시점까지의 정보를 바탕으로 k기 이후 즉, (n+k)기에 대한 예측치를 구하는 방법은 다음과 같음

$$y_t = e_t + \theta e_{n-1}$$

- 위 식에서 (n+1)기(k=1)의 $y$ 즉, $y_{n+1}$은 다음과 같이 나타낼 수 있음

$$y_{n+1} = e_{n+1} + \theta e_n$$

- 따라서 n시점까지의 정보를 이용하여 (n+1)기에 대한 예측치를 구하면 다음과 같음.

$$\hat{y}_{n+1} = \hat{\theta}\,\hat{e}_n \quad \text{또는} \quad E[y_{n+1} \mid I_t] = \theta e_n \leftarrow E[e_{n+1} \mid I_t = 0]$$

- 따라서, n시점까지의 정보를 이용하여 (n+1)기 예측치에 대한 예측오차(실제치-예측치)와 그 분산은 다음과 같음

- 예측오차 :
$$\begin{aligned} & y_{n+1} - E[y_{n+1} \mid I_t] \\ &= e_{n+1} + \theta e_n - \theta e_n \\ &= e_{n+1} \end{aligned}$$

- 예측오차 분산 :
$$\begin{aligned} & Var(y_{n+1} - E[y_{n+1} \mid I_t]) \\ &= Var(e_{n+1}) \\ &= \sigma_e^2 \end{aligned}$$

- 유사한 방법으로 n시점에서의 (n+2)기에 대한 예측치는 다음과 같음

$y_{n+2} = e_{n+2} + \theta e_{n+1}$이므로

$\hat{y}_{n+2} = 0$

또는 $E[y_{n+2} \mid I_t] = 0 \leftarrow E[e_{n+2} \mid I_t = 0]$ , $E[e_{n+1} \mid I_t = 0]$

- 따라서, n시점까지의 정보를 이용하여 (n+2)기 예측치에 대한 예측오차(실제치-예측치)와 그 분산은 다음과 같음

- 예측오차 : $y_{n+2} - E[y_{n+2} \mid I_t]$
$= e_{n+2} + \theta e_{n+1}$

- 예측오차 분산 : $Var(y_{n+2} - E[y_{n+2} \mid I_t])$
$= Var(e_{n+2} + \theta e_{n+1})$
$= (1 + \theta^2)\sigma_e^2$

```
 [1] -0.45537808  1.54818192 -0.04090282  0.19626161  1.63159555  0.29484781 -1.20124821 -0.55910317
 [9] -0.44443127  1.28446667  0.14442029  0.50975808 -0.02593756 -0.47022011  1.80123788  0.23737511
[17] -1.84351271  0.88530355 -0.70696988 -0.82844324 -0.24574157 -0.97497592 -0.62315286 -0.60133103
[25] -1.61681881  1.02198306 -0.10007430 -0.97207254  1.29012586  0.22424661 -0.20510301  0.86982934
[33]  0.77166096  0.77754365  0.60631904  0.51851340 -0.11318605 -0.25936838 -0.37215350 -0.64764980
[41] -0.14557427 -1.28244281  2.35854038  0.76435676 -0.96384485 -0.36364417 -0.43934215  0.82502157
[49] -0.22578665  0.37034654 -0.15055853  0.05139303  1.30485660 -0.37001480  1.65476546 -1.86315759
[57]  1.03382926 -0.29026166  0.50500577  0.13557319 -0.37484432 -0.35727087 -0.95417104 -0.97698787
[65]  0.38328365  0.34675249  0.06538340  0.90569575  1.93336129 -0.69138813 -2.09226034  1.16820729
[73] -0.97006211 -0.39582339  0.90561827 -0.33945674 -1.14043261  0.29262807 -0.24661173  0.10488639
[81]  0.31113815 -0.36968903  0.69549695 -0.34842780  0.45727622  0.95758992  0.38480683 -0.34952410
[89]  1.21184596  0.78619766  0.56282670  0.15135955 -0.59665121  1.42534442 -0.83841687  2.44748448
[97]  1.03422583 -0.08131154 -1.10768233 -0.50673436
```

**(점 예측치)**

**-n시점에서의 (n+1)기에 대한 예측치**

$$\hat{y}_{t+1} = \hat{\theta}\hat{e}_n = 0.7399(-0.5067) = -0.3749$$

**-n시점에서의 (n+2)기에 대한 예측치**

$$\hat{y}_{t+2} = 0$$

**(구간예측)**

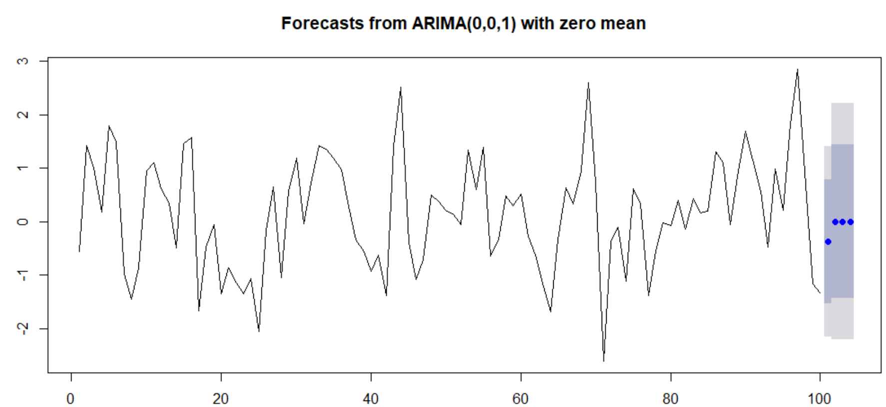
**-n시점에서의 (n+1)기에 대한 95% 구간예측**

$$-0.3749 \pm (1.96)\sqrt{0.8194} = [-2.1491,\ 1.3992]$$

**-n시점에서의 (n+2)기에 대한 95% 구간예측**

$$0 \pm (1.96)\sqrt{(1+0.7399^2)(0.8194)} = [-2.207,\ 2.207]$$

```
> predict(ma1,n.ahead=4)
$pred
Time Series:
Start = 101
End = 104
Frequency = 1
[1] -0.3749101  0.0000000  0.0000000  0.0000000

$se
Time Series:
Start = 101
End = 104
Frequency = 1
[1] 0.9052254 1.1260459 1.1260459 1.1260459
> (fma1<-forecast(ma1,h=4))
    Point Forecast     Lo 80     Hi 80     Lo 95    Hi 95
101     -0.3749101 -1.535003 0.7851829 -2.149119 1.399299
102      0.0000000 -1.443086 1.4430859 -2.207009 2.207009
103      0.0000000 -1.443086 1.4430859 -2.207009 2.207009
104      0.0000000 -1.443086 1.4430859 -2.207009 2.207009
```



## (1) 모형

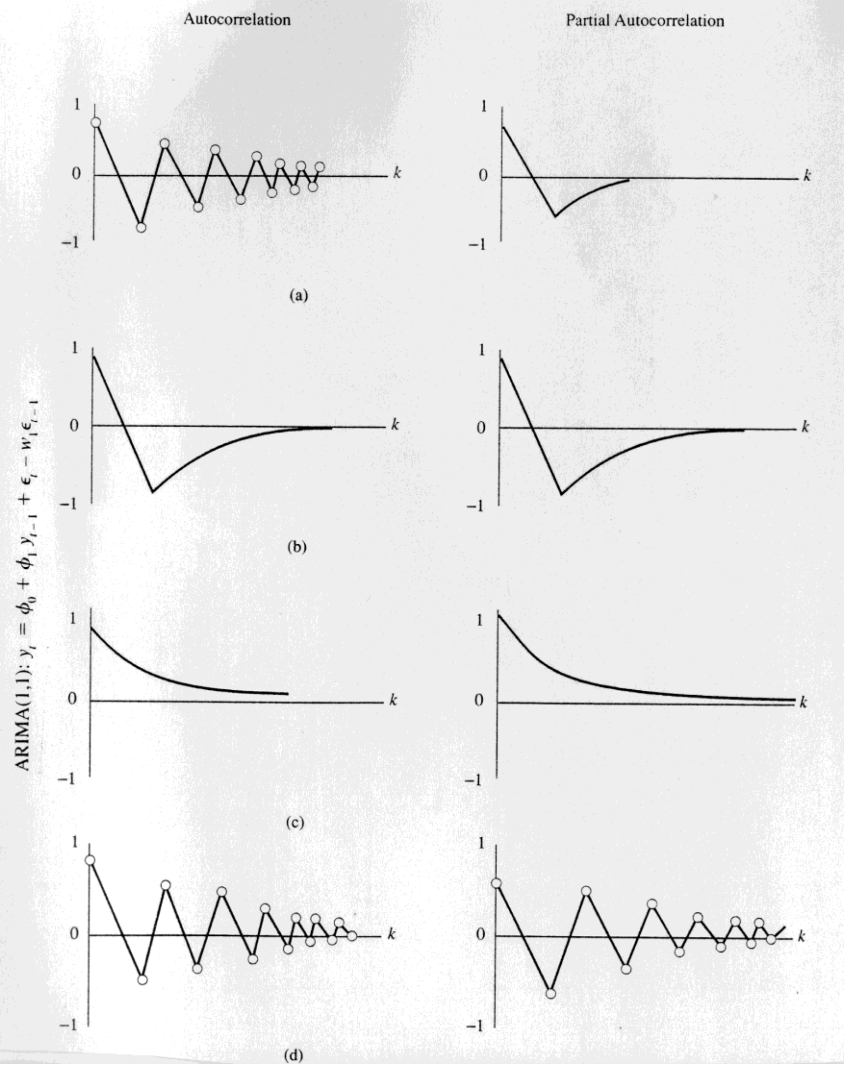
ARMA(1,1) $y_t = \mu + \phi\, y_{t-1} + e_t - \theta e_{t-1}$ , $e_t \sim iidN(0, \sigma_e^2)$

ARMA(p,q) $y_t = \mu + \phi_1\, y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$ , $e_t \sim iidN(0, \sigma_e^2)$

## (2) ARMA(1,1)

① 식별

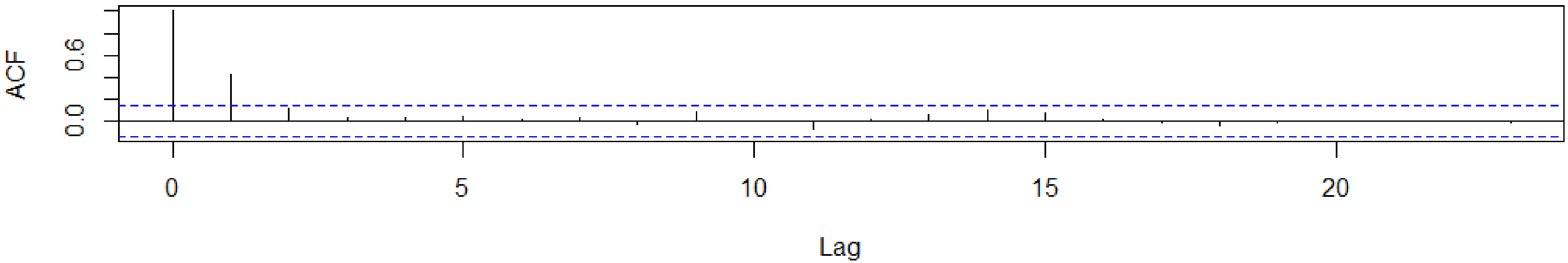
- ACF는 시차 q(=1)까지는 0과 현저히 다르고, 그 이후는 AR(p)처럼 점차적으로 소멸함
- PACF는 시차 p(=1)까지는 0과 현저히 다르고, 그 이후는 MA(q)처럼 점차적으로 소멸함



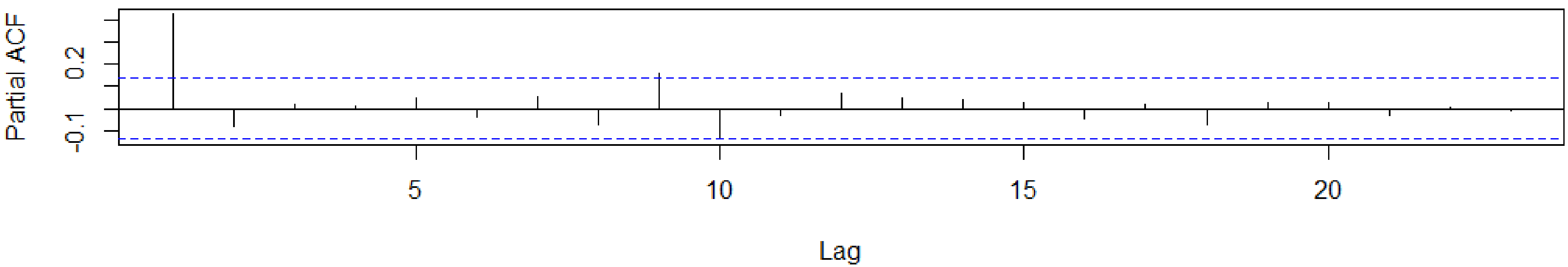
주요확률과정의 acf와 pacf (요약)

| 확률과정 | acf | pacf |
|---|---|---|
| AR | 점차적으로 0으로 수렴함<br>(tail off) | 시차 p까지만 0과 현저히 다르고 이후 0으로 갑자기 이동<br>(즉, 마지막 스파이크의 시차길이는 AR의 차수와 동일)<br>(cuts off after lag p) |
| MA | 시차 q까지만 0과 현저히 다르고 이후 0으로 갑자기 이동<br>(즉, 마지막 스파이크의 시차길이는 MA의 차수와 동일)<br>(cuts off after lag q) | 점차적으로 0으로 수렴함<br>(tail off) |
| ARMA | 시차 q까지는 0과 현저히 다르고,<br>그 이후는 AR(p)처럼 점차적으로 소멸함 | 시차 p까지는 0과 현저히 다르고,<br>그 이후는 MA(q)처럼 점차적으로 소멸함 |
| white noise | 모든 시차에 대해 0 | 모든 시차에 대해 0 |

**Y(t) = 0.3Y(t-1)+e(t)+0.2e(t-1)**

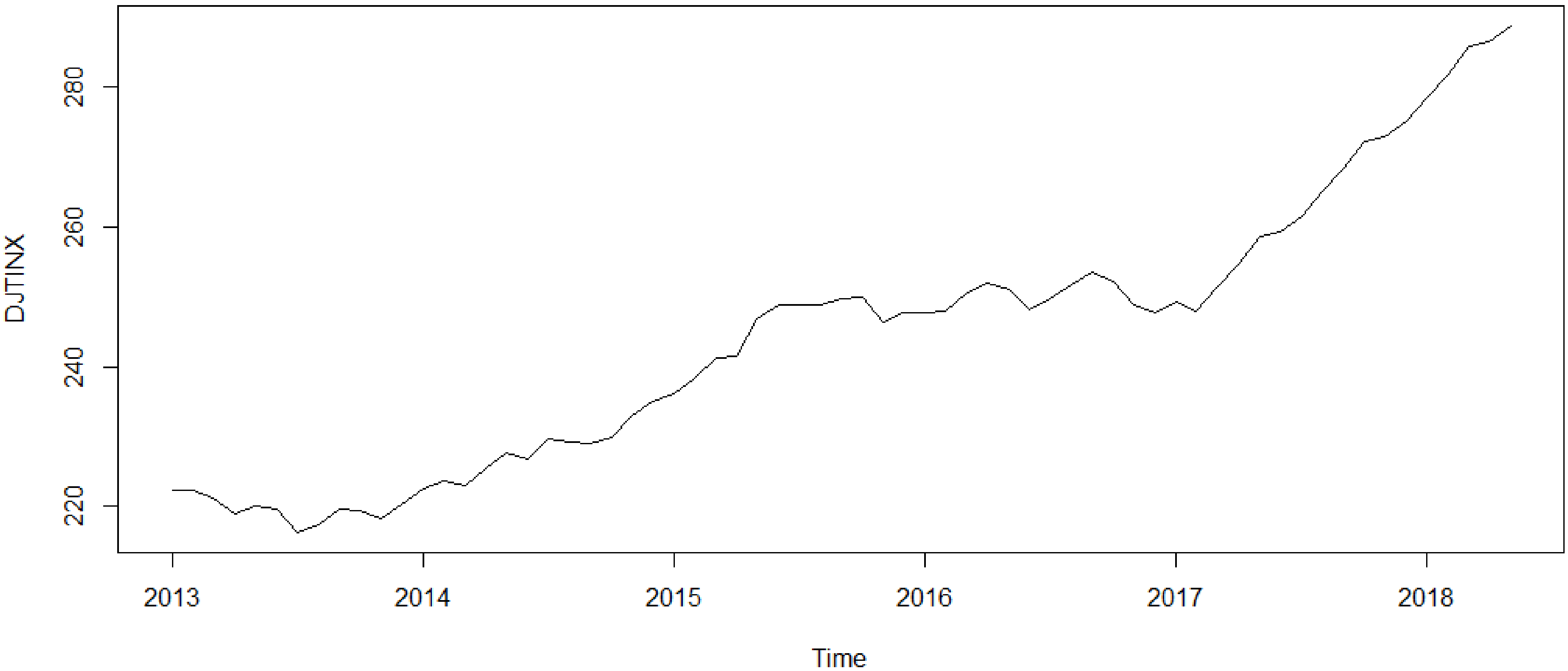



**Y(t) = 0.3Y(t-1)+e(t)+0.2e(t-1)**

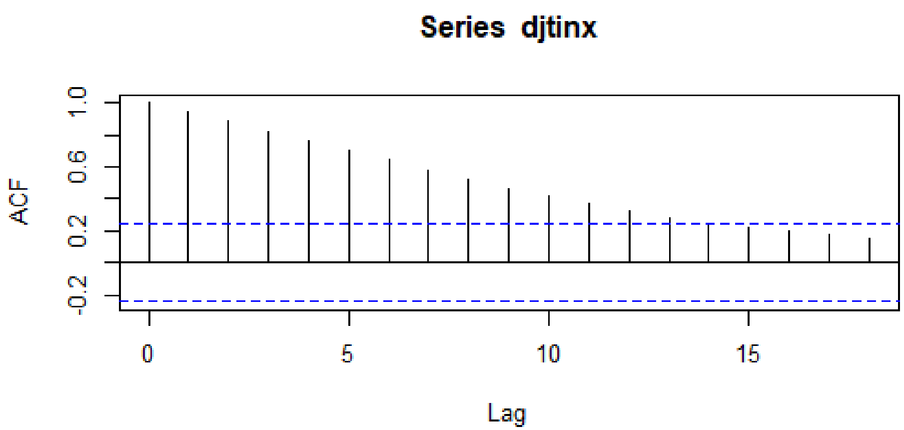


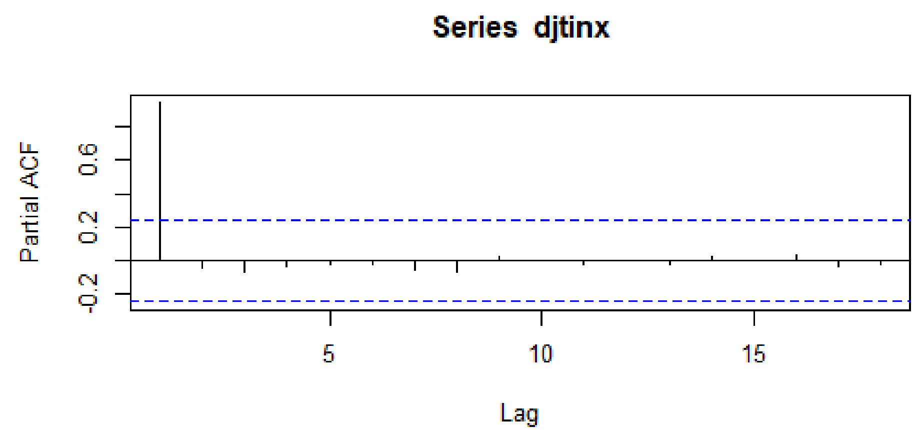
# (3) 사례분석(ARMA(1,1))

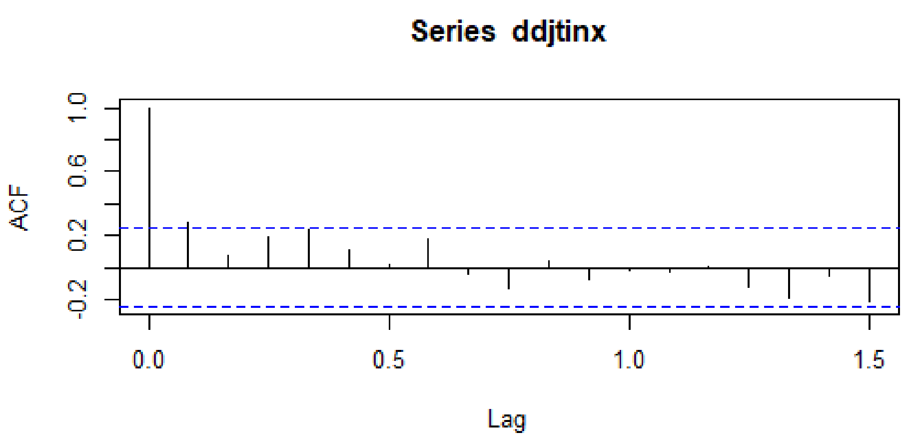
① 데이터

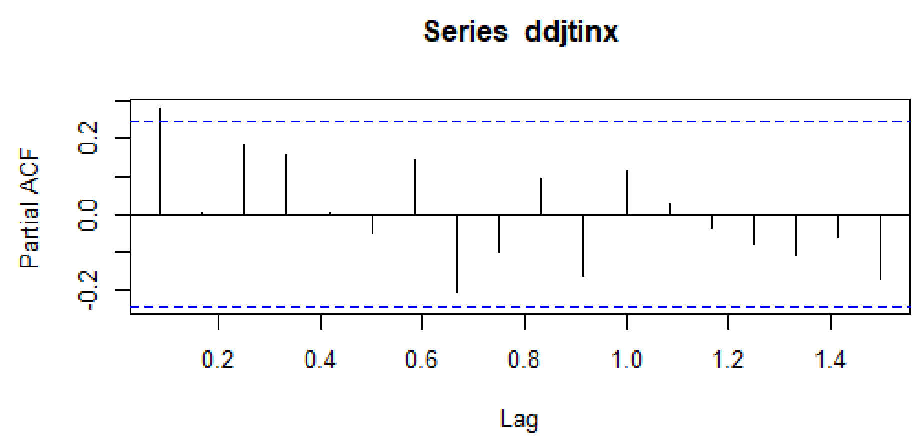


## ② 식별


Series djtinx
ACF
-0.2
0.2
0.6
1.0
0
5
10
15
Lag


Series djtinx
Partial ACF
-0.2
0.2
0.6
5
10
15
Lag

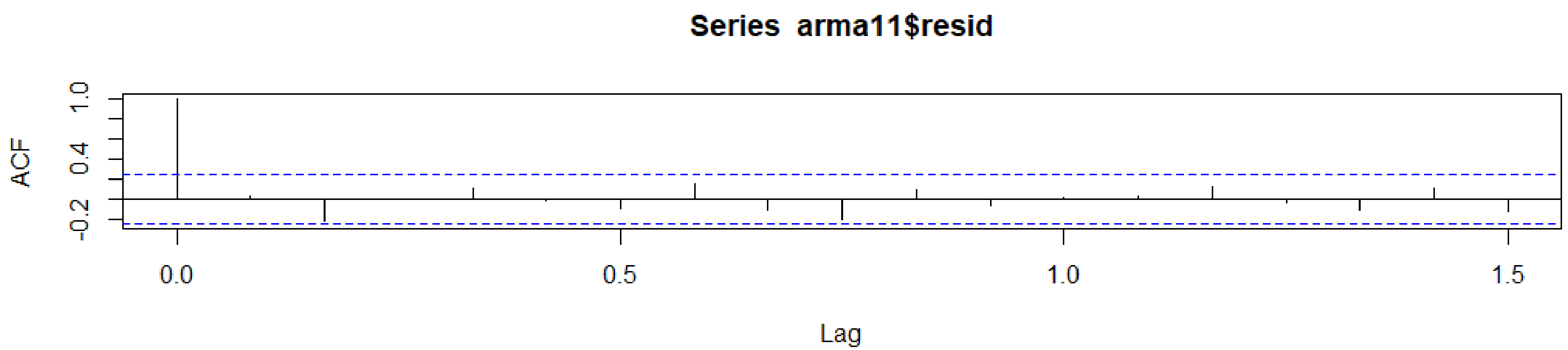

Series ddjtinx
ACF
-0.2
0.2
0.6
1.0
0.0
0.5
1.0
1.5
Lag

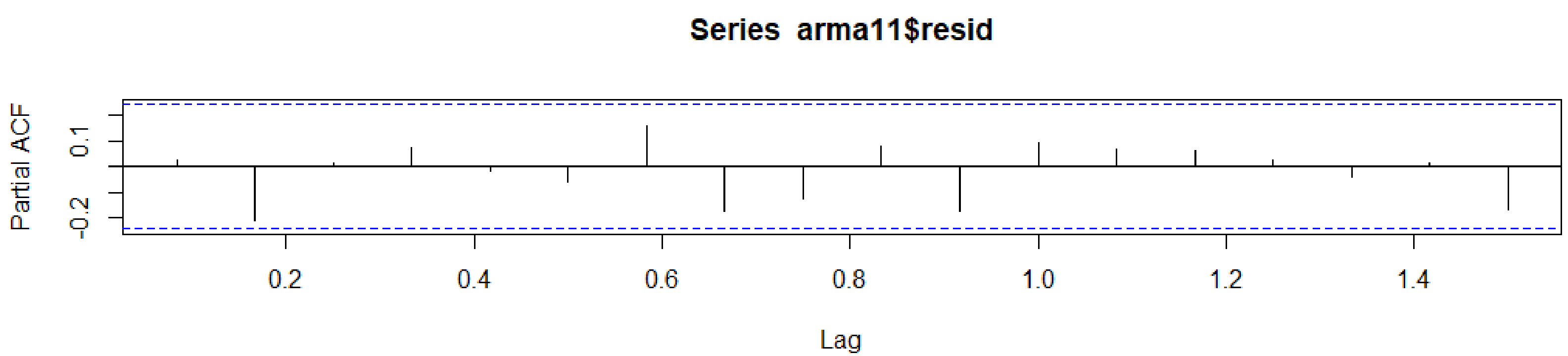

Series ddjtinx
Partial ACF
-0.2
0.0
0.2
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Lag

### ③ 추정

```
> (arma11<-arima(ddjtinx, order=c(1,0,1), include.mean=F))

Call:
arima(x = ddjtinx, order = c(1, 0, 1), include.mean = F)

Coefficients:
         ar1      ma1
      0.9482  -0.7407
s.e.  0.0600   0.1221

sigma^2 estimated as 3.456:  log likelihood = -130.82,  aic = 267.65
```

### ④ 모형의 검토



⑤ 예측

– ARMA(1,1)의 경우를 예를 들어 n시점까지의 정보를 바탕으로 k기 이후 즉, (n+k)기에 대한 예측치를 구하는 방법은 다음과 같음

$$y_t = \mu + \phi y_{t-1} + \theta e_{t-1} + e_t$$

- 위 식에서 (n+1)기의 $y$ 즉, $y_{n+1}$은 다음과 같이 나타낼 수 있음

$$y_{n+1} = \mu + \phi y_n + \theta e_n + e_{n+1}$$

- 따라서 n시점까지의 정보를 이용하여 (n+1)기에 대한 예측치를 구하면 다음과 같음.

$$\hat{y}_{n+1} = \hat{\mu} + \hat{\phi} y_n + \hat{\theta} e_n \quad \text{또는} \quad E[y_{n+1} \mid I_t] = \mu + \phi y_n + \theta e_n \leftarrow E[e_{n+1} \mid I_t = 0]$$

- 따라서, n시점까지의 정보를 이용하여 (n+1)기 예측치에 대한 예측오차(실제치-예측치)와 그 분산은 다음과 같음

- 예측오차 :
$$\begin{aligned} & y_{n+1} - E[y_{n+1} \mid I_t] \qquad | \quad (*) \\ &= \mu + \phi y_n + \theta e_n + e_{n+1} - (\mu + \phi y_n + \theta e_n) \\ &= e_{n+1} \end{aligned}$$

- 예측오차 분산 :
$$\begin{aligned} & Var(y_{n+1} - E[y_{n+1} \mid I_t]) \\ &= Var(e_{n+1}) \\ &= \sigma_e^2 \end{aligned}$$

- 유사한 방법으로 (n+2)기의 $y$ 즉, $y_{n+2}$은 다음과 같이 나타낼 수 있음

$$y_{n+2} = \mu + \phi y_{n+1} + \theta e_{n+1} + e_{n+2}$$

- 따라서 n시점에서의 (n+2)기에 대한 예측치는 다음과 같음

$$\begin{aligned}\hat{y}_{n+2} &= \hat{\mu} + \hat{\phi}\hat{y}_{n+1} \\ &= \hat{\mu} + \hat{\phi}(\hat{\mu} + \hat{\phi} y_n + \hat{\theta} e_{n)} \\ &= (\hat{\phi} + 1)\hat{\mu} + \widehat{\phi^2} y_n + \hat{\phi}\hat{\theta} e_n\end{aligned}$$

또는 $E[y_{n+2} \mid I_t] = \mu + \phi E[y_{n+1} \mid I_t] \leftarrow E[\theta e_{n+1} \mid I_t = 0],\ E[e_{n+2} \mid I_t = 0]$

- 따라서, n시점까지의 정보를 이용하여 (n+2)기 예측치에 대한 예측오차(실제치-예측치)와 그 분산은 다음과 같음

- 예측오차 : $y_{n+2} - E[y_{n+2} \mid I_t]$

$$= \mu + \phi y_{n+1} + \theta e_{n+1} + e_{n+2} - \mu - \phi E[y_{n+1} \mid I_t]$$

$$= \phi e_{n+1} + \theta e_{n+1} + e_{n+2}$$

$$\because \ \phi y_{n+1} - \phi E[y_{n+1} \mid I_t] = \phi\,(y_{n+1} - E[y_{n+1} \mid I_t]) = \phi e_{n+1} \text{ from(※)}$$

- 예측오차 분산 : $Var(y_{n+2} - E[y_{n+2} \mid I_t])$

$$= Var(\phi e_{n+1} + \theta e_{n+1} + e_{n+2})$$

$$= (\phi^2 + \theta^2 + 1 + 2\phi\theta)\sigma_e^2$$

- 일반적으로 n시점에서 h기 이후 즉, (n+h)기에 대한 예측치는 다음과 같음

$$\hat{y}_{n+h} = (\hat{\phi}^{h-1} + \hat{\phi}^{h-2} + \ldots + 1)\hat{\mu} + \hat{\phi}^{h} y_n + \hat{\phi}^{h-1}\hat{\theta} e_n$$

(점 예측치)

-n시점에서의 (n+1)기에 대한 예측치

$$\Delta\hat{y}_{t+1} = \hat{\phi}\Delta y_t + \hat{\theta}e_t$$

$$\begin{aligned}\Delta y_{(2018:6)} &= (0.9482)\Delta y_{(2018:5)} + (-0.7407)e_{2018:5} \\ &= (0.9482*2.24) + (-0.7407*0.29193) \\ &= 2.124 - 0.2162 \\ &= 1.9077\end{aligned}$$

-n시점에서의 (n+2)기에 대한 예측치

$$\Delta\hat{y}_{t+2} = \hat{\phi}\Delta\hat{y}_{t+1}$$

$$\begin{aligned}\Delta y_{(2018:7)} &= (0.9482)\Delta y_{(2018:6)} \\ &= (0.9482*1.9077) \\ &= 1.8088\end{aligned}$$

(구간예측)

-n시점에서의 (n+1)기에 대한 95% 구간예측

$$1.9077 \pm (1.96)\sqrt{3.456} = [-1.735737,\ 5.551083]$$

-n시점에서의 (n+2)기에 대한 95% 구간예측

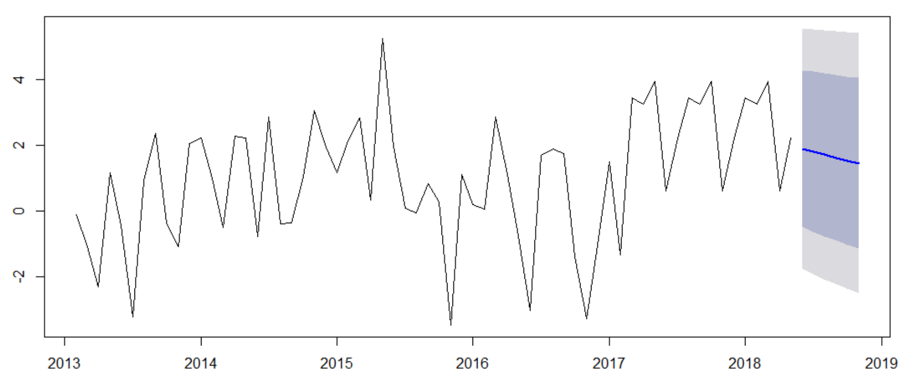
$$1.8088 \pm (1.96)\sqrt{((0.9482^2 + (-0.7407)^2 + 1 - 2*0.9482*0.74407)3.456}$$
$$= [-1.912163,\ 5.529787]$$

```
> (ddjtinx<-djtinx.ts-lag(djtinx.ts,k=-1))
       Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct   Nov   Dec
2013       -0.10 -1.07 -2.29  1.17 -0.44 -3.21  0.93  2.36 -0.37 -1.07  2.05
2014  2.24  1.02 -0.49  2.29  2.24 -0.78  2.87 -0.39 -0.34  1.03  3.06  1.95
2015  1.17  2.14  2.83  0.34  5.26  1.99  0.10 -0.05  0.83  0.29 -3.45  1.12
2016  0.19  0.05  2.87  1.12 -0.73 -3.02  1.71  1.90  1.75 -1.37 -3.26 -1.02
2017  1.51 -1.32  3.46  3.26  3.95  0.63  2.24  3.46  3.26  3.95  0.63  2.24
2018  3.46  3.26  3.95  0.63  2.24
```

```
> res
              Jan          Feb          Mar          Apr          May          Jun          Jul          Aug
2013              -0.083734825 -0.952026376 -1.858598917  1.976374839 -0.112227150 -2.859156128  1.861863574
2014  1.967708905  0.353391736 -1.195338364  1.869169358  1.453183120 -1.827491752  2.255888563 -1.440252482
2015 -0.003132137  1.028312424  1.562609468 -1.185859240  4.059209999  0.009394083 -1.779914250 -1.463265888
2016  0.072756567 -0.076260210  2.766102384  0.447685298 -1.460341540 -3.409559448  2.047908903  1.795578536
2017  2.088960102 -1.204377869  3.819466887  2.808526658  2.939321382 -0.938036467  0.947810084  2.038160785
2018  1.640915768  1.194793452  1.743969648 -1.823477606  0.291931158
```

```
> (f3<-forecast(arma11, h=6))
         Point Forecast      Lo 80    Hi 80     Lo 95    Hi 95
Jun 2018       1.907673 -0.4746248 4.289971 -1.735737 5.551083
Jul 2018       1.808812 -0.6242027 4.241827 -1.912163 5.529787
Aug 2018       1.715075 -0.7626509 4.192800 -2.074279 5.504429
Sep 2018       1.626195 -0.8910495 4.143439 -2.223598 5.475988
Oct 2018       1.541921 -1.0103300 4.094172 -2.361410 5.445252
Nov 2018       1.462014 -1.1213041 4.045333 -2.488830 5.412859
```

Forecasts from ARIMA(1,0,1) with zero mean




```
> predict(arma11,n.ahead=4)
$pred
          Jun      Jul      Aug      Sep
2018 1.907673 1.808812 1.715075 1.626195

$se
          Jun      Jul      Aug      Sep
2018 1.858917 1.898492 1.933379 1.964216
```