방한관광객수 예측 모델

관광객의 방한 의도에 영향을 미치는건
경제냐? 아님 우호도냐?

데이터 설명

변수명	설명
PAX	일본인 방한관광객수 (1987 ~ 2017,연간)
Cur	원-엔 환율 (1987 ~ 2017, 연간)
Fri_ab	국가 우호도(1987 ~ 2017,연간)

- 목적

1) 방한관광객수에 대해 각각
(1) 환율에 대한 CCF를 확인하고
(2) 국가 우호도에 대한 CCF를 확인한다.

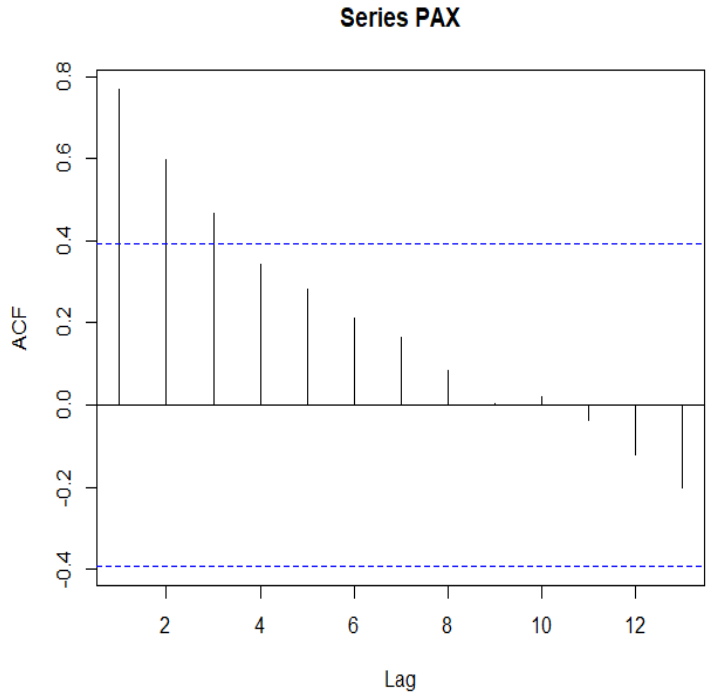
2) 각각의 변수와 그 복합 모델에 대하여
(1) 모델을 적합하고, 그 예측력을 확인한다.

사전 백색화

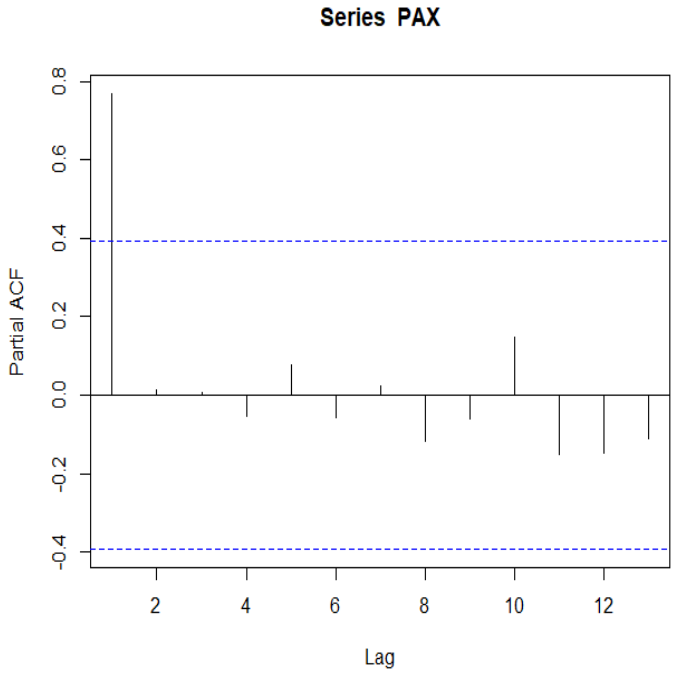
- ARIMA 모형 적합

- PAX(방한관광객수)

1) 차분전




```
Type 1: no drift no trend
     lag  ADF p.value
[1,]   0 1.79   0.978
[2,]   1 1.88   0.981
[3,]   2 1.92   0.982
Type 2: with drift no trend
     lag    ADF p.value
[1,]   0 -0.252   0.916
[2,]   1  0.209   0.965
[3,]   2  0.418   0.978
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -2.37   0.415
[2,]   1 -2.19   0.483
[3,]   2 -2.12   0.507
```

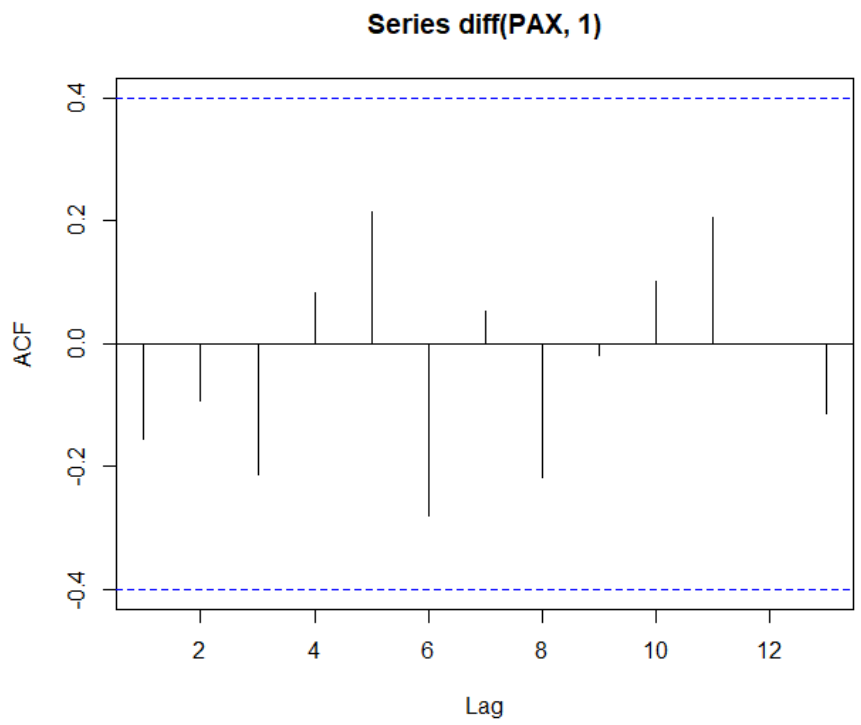


사전 백색화

- ARIMAX 모형 적합

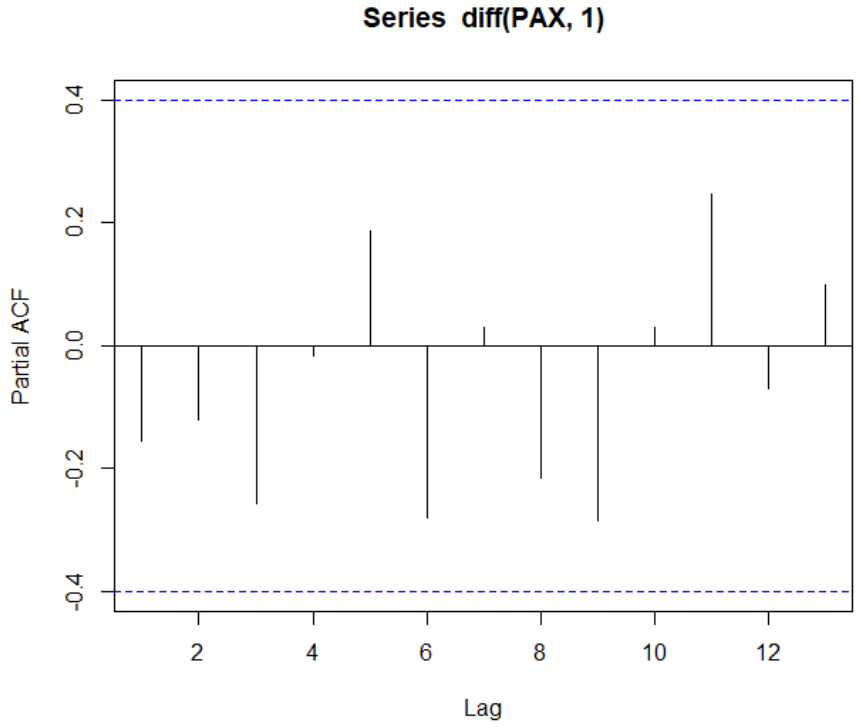
- PAX(방한관광객수)

1) 차분후




```
Type 1: no drift no trend
     lag   ADF p.value
[1,]   0 -4.77  0.0100
[2,]   1 -3.03  0.0100
[3,]   2 -2.50  0.0158
Type 2: with drift no trend
     lag   ADF p.value
[1,]   0 -5.38  0.0100
[2,]   1 -3.72  0.0111
[3,]   2 -3.48  0.0196
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -5.45  0.0100
[2,]   1 -3.86  0.0313
[3,]   2 -3.72  0.0411
```

<ADF 테스트>

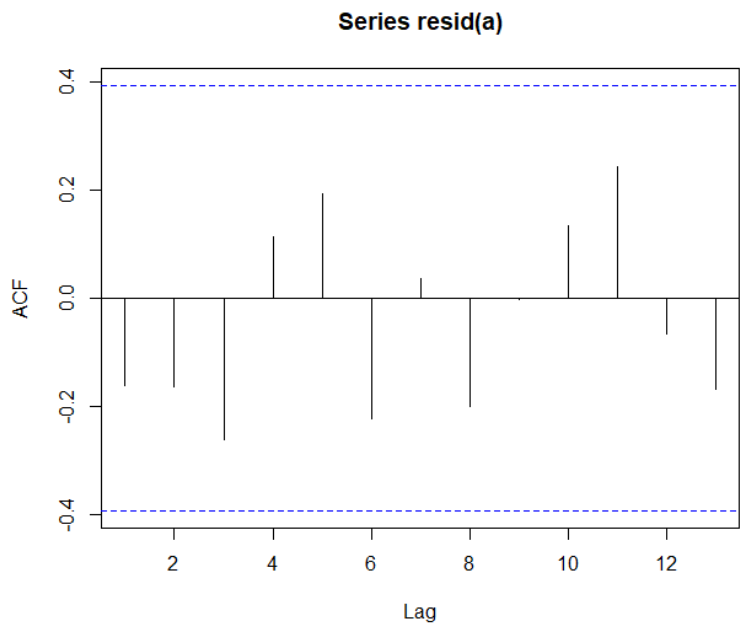


사전 백색화

- ARIMA 모형 적합

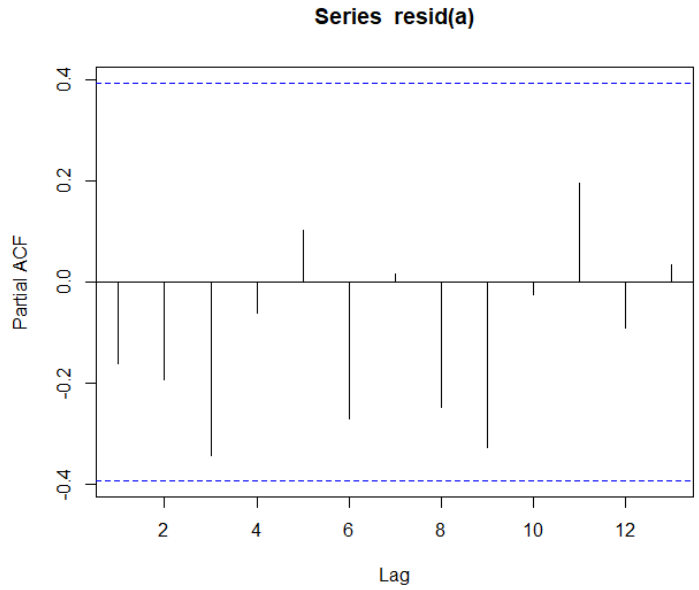
-PAX(방한관광객수)

3) ARIMA 모형 적합(잔차 검정)




```
data:  resid(a)
X-squared = 20.311, df = 19, p-value = 0.3761
```

<BOX-Ljung>



그래인저 인과성 테스트

- 방한관광객수, 우호도

```
Model 1: PAX_1 ~ Lags(PAX_1, 1:4) + Lags(fri_ab_diff, 1:4)
Model 2: PAX_1 ~ Lags(PAX_1, 1:4)
  Res.Df Df      F Pr(>F)
1     11
2     15 -4 1.6227 0.2373
```

```
Model 1: fri_ab_diff ~ Lags(fri_ab_diff, 1:4) + Lags(PAX_1, 1:4)
Model 2: fri_ab_diff ~ Lags(fri_ab_diff, 1:4)
  Res.Df Df      F Pr(>F)
1     11
2     15 -4 1.6892 0.2221
> |
```

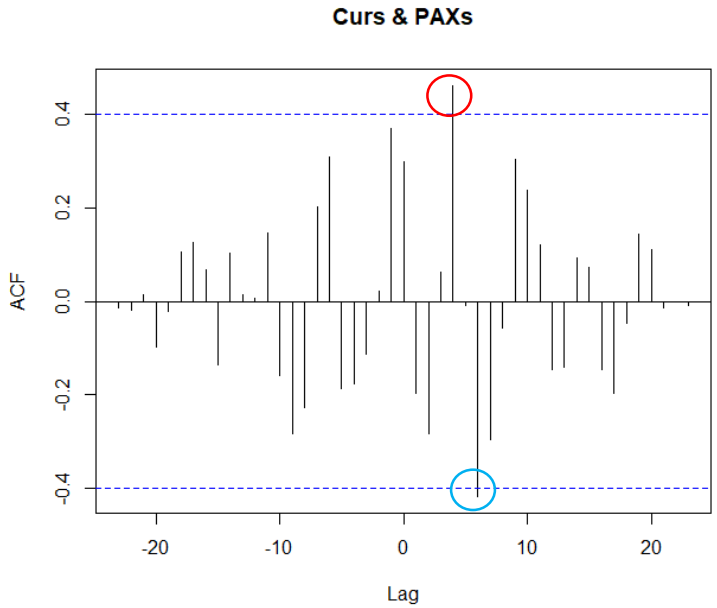
방한관광객수와 우호도는 그래인저 인과성이 존재하지 않는다.

```
Model 1: PAX_1 ~ Lags(PAX_1, 1:4) + Lags(Cur_diff, 1:4)
Model 2: PAX_1 ~ Lags(PAX_1, 1:4)
  Res.Df Df      F Pr(>F)
1     11
2     15 -4 1.0188 0.4394
```

```
Model 1: Cur_diff ~ Lags(Cur_diff, 1:4) + Lags(PAX_1, 1:4)
Model 2: Cur_diff ~ Lags(Cur_diff, 1:4)
  Res.Df Df      F Pr(>F)
1     11
2     15 -4 1.4822 0.2732
> |
```

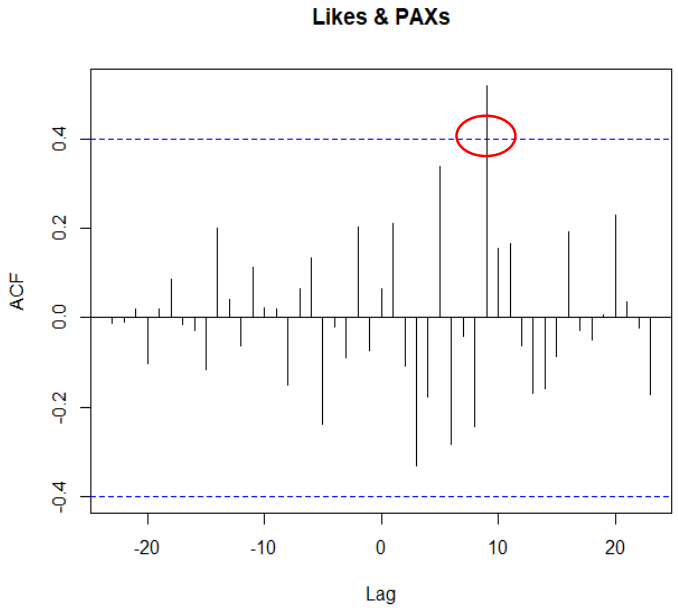
방한관광객수와 환율 그래인저 인과성이 존재하지 않는다.

CCF 구조 파악



지연모수 : 4
투입모수 : 0
산출모수 : 2(only)

$$Y_t = \frac{w_0 B_0}{1 - \delta_2} X_{t-4} + \epsilon$$



지연모수 : 8
투입모수 : 0
산출모수 : 0

$$Y_t = w_0 B_8 X_t + \epsilon$$

모형 가적합

- 방한관광객수, 환율

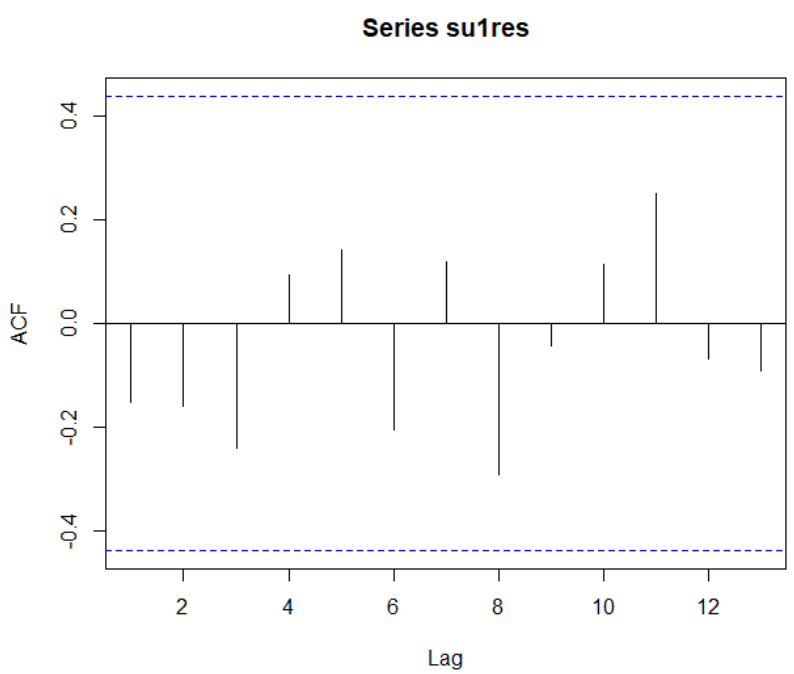
- 모형 가적합

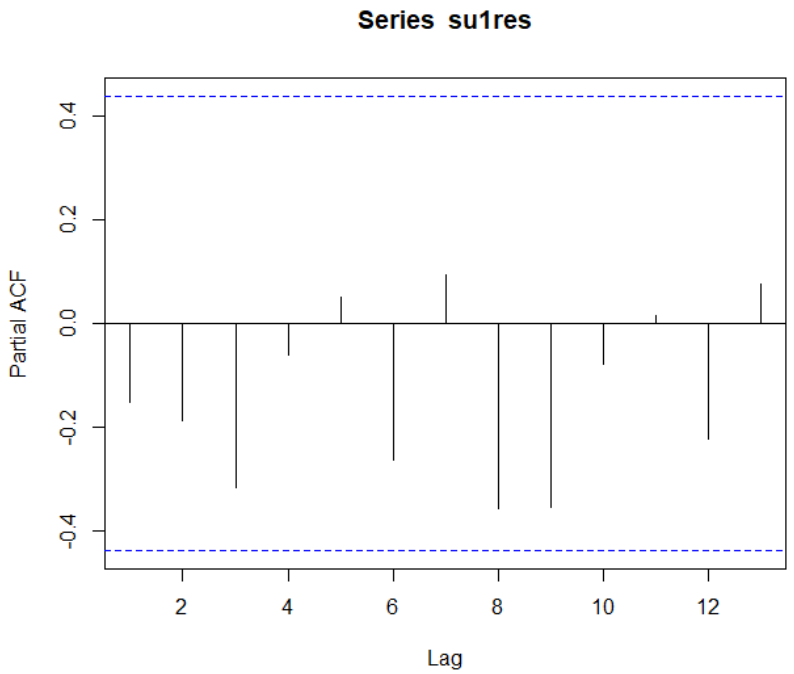
```
Call:
arimax(x = su1[, 1], order = c(1, 0, 0), include.mean = FALSE, fixed = c(NA,
    0, NA, NA), xtransf = su1[, 2], transfer = list(c(2, 0)))

Coefficients:
          ar1  T1-AR1  T1-AR2   T1-MA0
      -0.0699       0  0.2648  -0.0523
s.e.   0.2234       0  1.2083   0.2191

sigma^2 estimated as 0.01581:  log likelihood = 13.09,  aic = -20.18
```

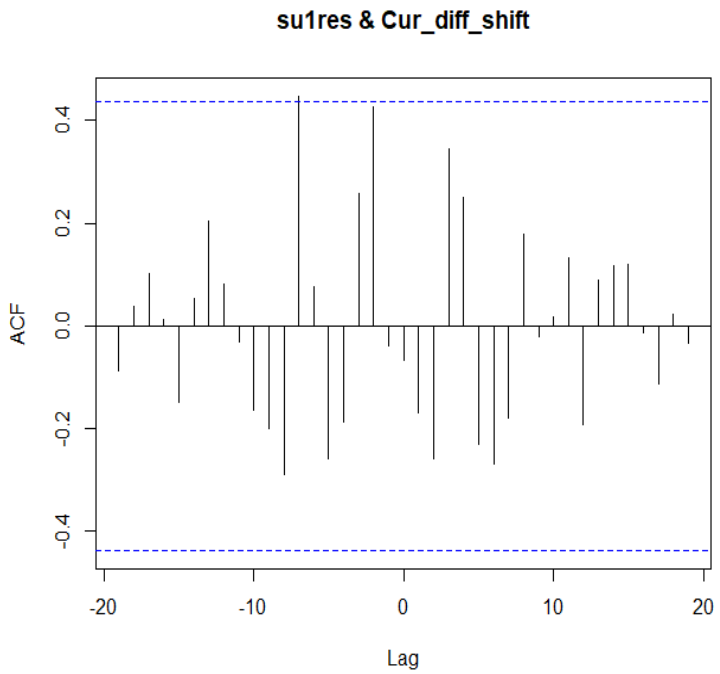
모형 가적합

- 잔차의 BOX-test 및 acf,pacf

```
        Box-Ljung test

data:  su1res
X-squared = 0.54489, df = 1, p-value = 0.4604
```

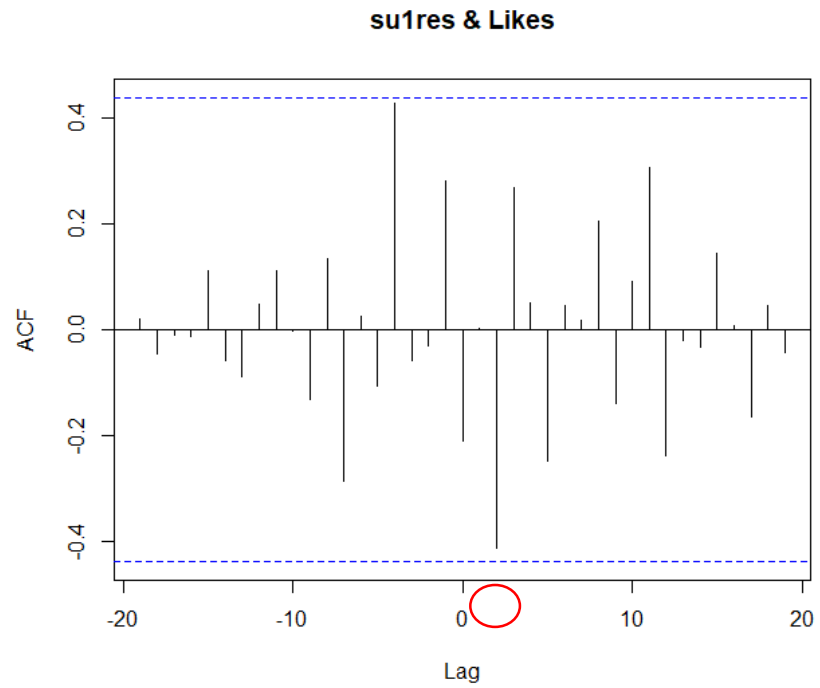






모형 가적합

- 추가 투입변수(우호도) 적합 가능성 탐색



지연모수 2
투입모수 0
산출모수 0

$$Y_t = \frac{w_0 B_0}{1 - \delta_2} X_{1,t-4} + w_1 B_0 X_{2,t-2} + \epsilon$$

모형 가적합

- 추가 투입변수(우호도) 적합 가능성 탐색

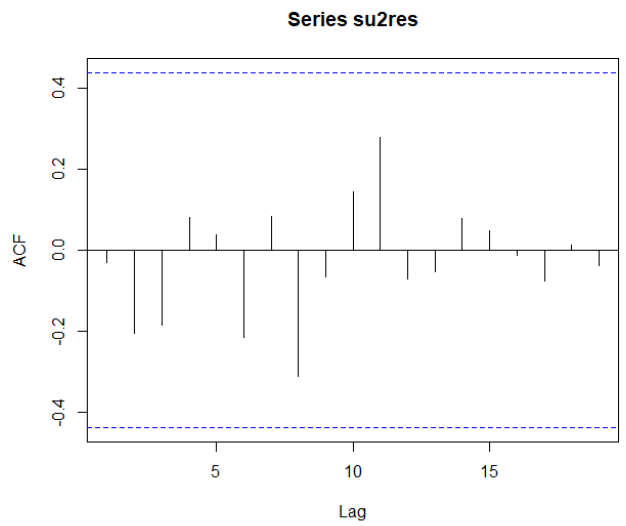
```
Call:
arimax(x = su1[, 1], order = c(1, 0, 0), fixed = c(NA, NA, 0, NA, NA, NA), xtransf = data.frame(su2[,
    2], su2[, 3]), transfer = list(c(2, 0), c(0, 0)))

Coefficients:
          ar1  intercept  su2...2.-AR1  su2...2.-AR2  su2...2.-MA0  su2...3.-MA0
      -0.2268     0.0648             0        0.7258       -0.2502       -0.0018
s.e.   0.2690     0.0346             0        0.3737        0.2021        0.0117

sigma^2 estimated as 0.01273:  log likelihood = 15.24,  aic = -20.47
> |
```

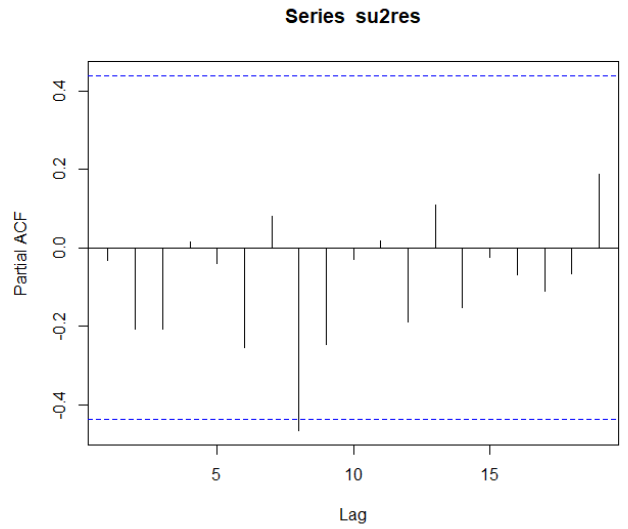
모형 가적합

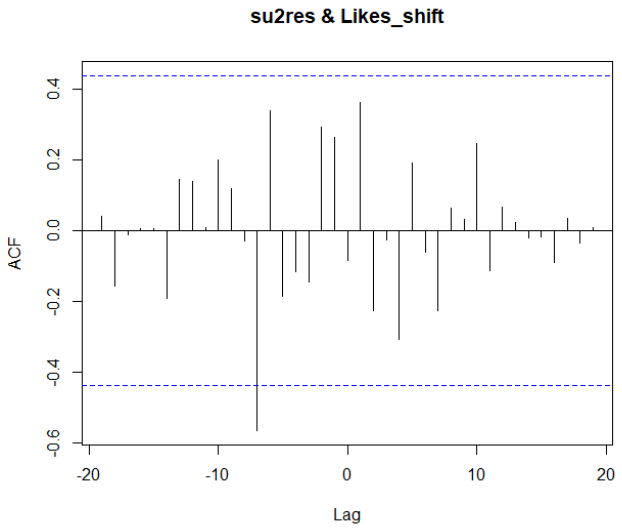
- 추가 투입변수(우호도) 적합 가능성 탐색

- 잔차의 BOX-test, acf, pacf

```
        Box-Ljung test

data:  su2res
X-squared = 0.02234, df = 1, p-value = 0.8812
```







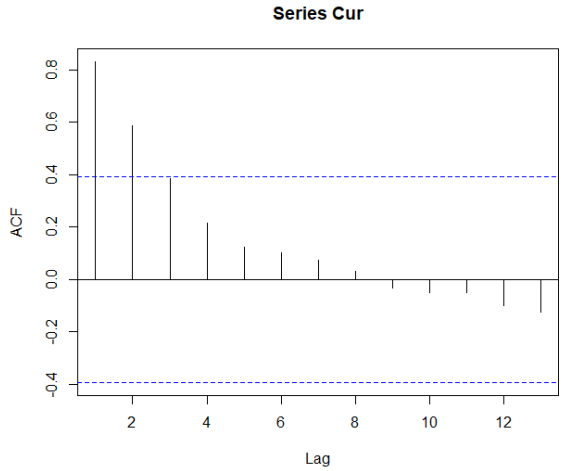
모형 테스트

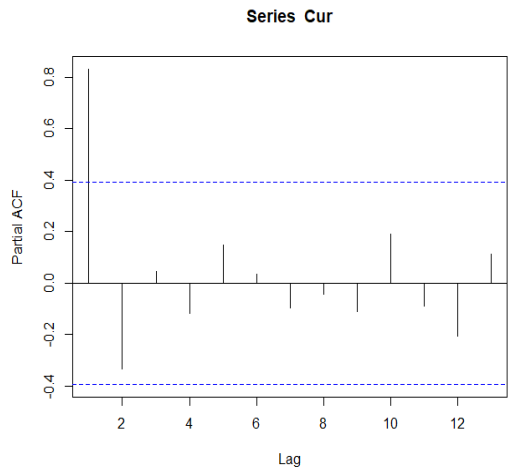
- ARIMA 모형 적합

- Cur(환율)

1) Adf 검정 및 acf,pacf

```
Type 1: no drift no trend
     lag  ADF p.value
[1,]   0 1.43   0.957
[2,]   1 1.36   0.952
[3,]   2 1.70   0.975
Type 2: with drift no trend
     lag    ADF p.value
[1,]   0 -0.868   0.728
[2,]   1 -1.858   0.382
[3,]   2 -1.413   0.540
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -1.98   0.566
[2,]   1 -2.99   0.190
[3,]   2 -2.18   0.484
```





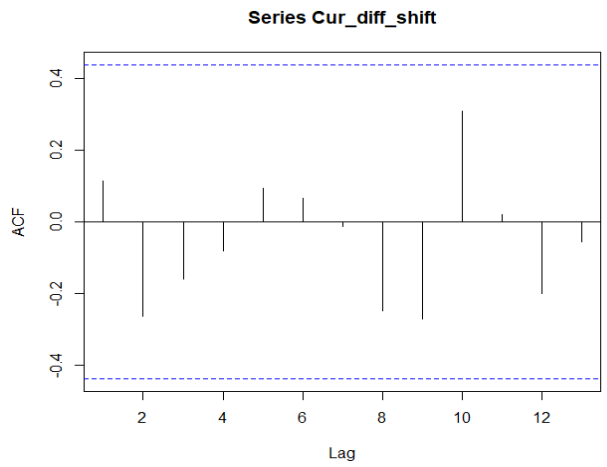
모형테스트

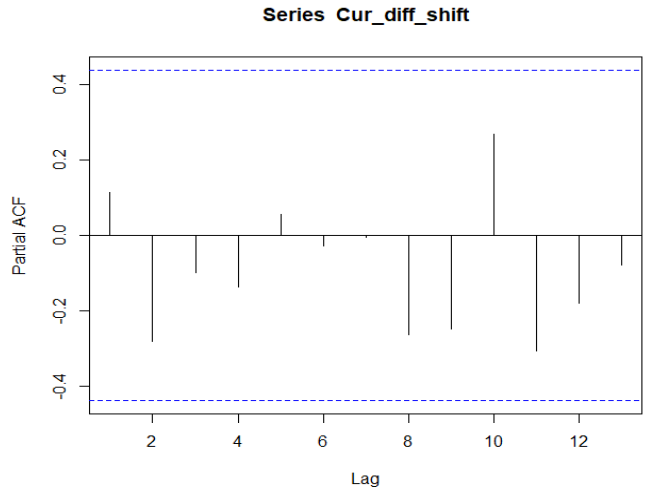
- ARIMA 모형 적합

- CUR(환율)

2) ADF 검정 및 acf,pacf(차분후)

```
Type 1: no drift no trend
     lag   ADF p.value
[1,]   0 -2.97  0.0100
[2,]   1 -2.98  0.0100
[3,]   2 -2.17  0.0325
Type 2: with drift no trend
     lag   ADF p.value
[1,]   0 -3.20  0.0347
[2,]   1 -3.36  0.0239
[3,]   2 -2.59  0.1171
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -3.20  0.1139
[2,]   1 -3.50  0.0632
[3,]   2 -3.12  0.1418
----
```



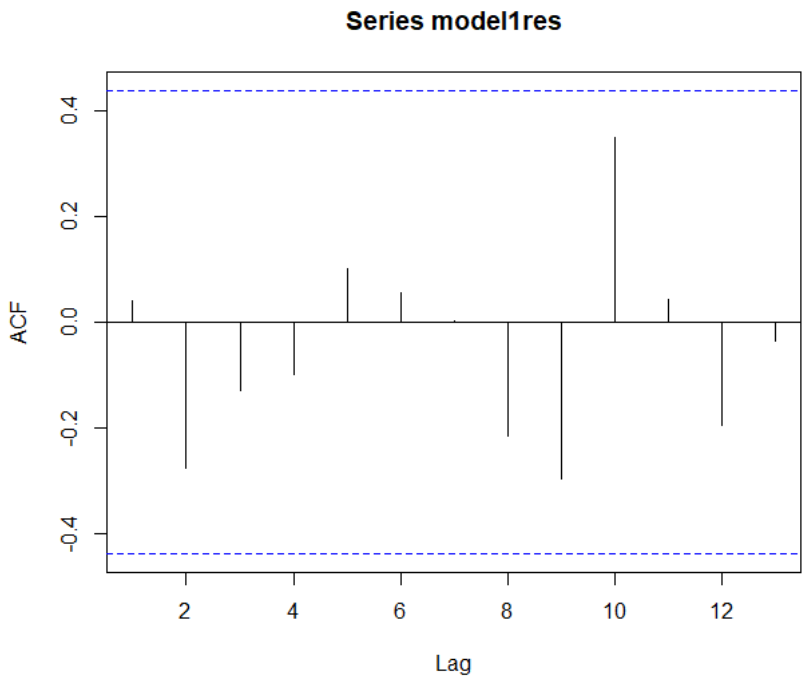


모형 테스트

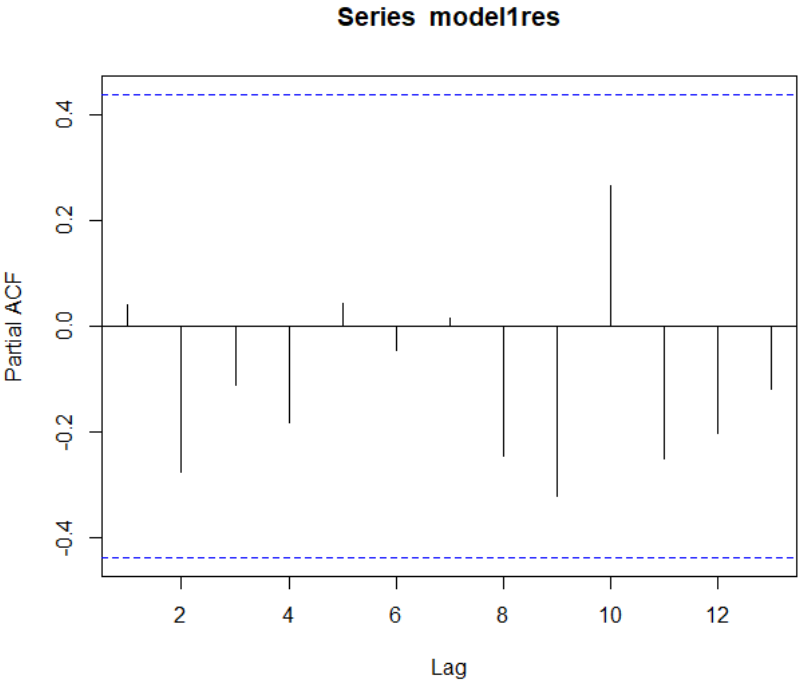
- ARIMA 모형 적합

- Cur(환율)

3) 임의로 ARIMA(1,1,0)을 적용한다.




```
        Box-Ljung test

data:  model1res
X-squared = 0.035317, df = 1, p-value = 0.8509
```



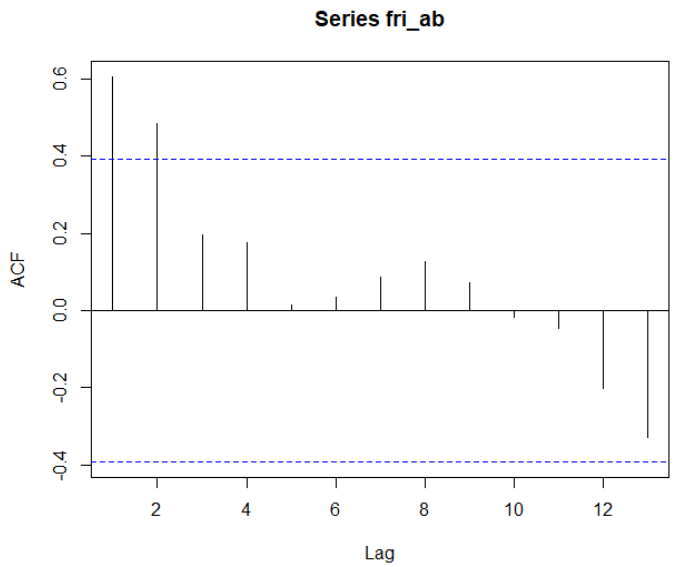
모형 테스트

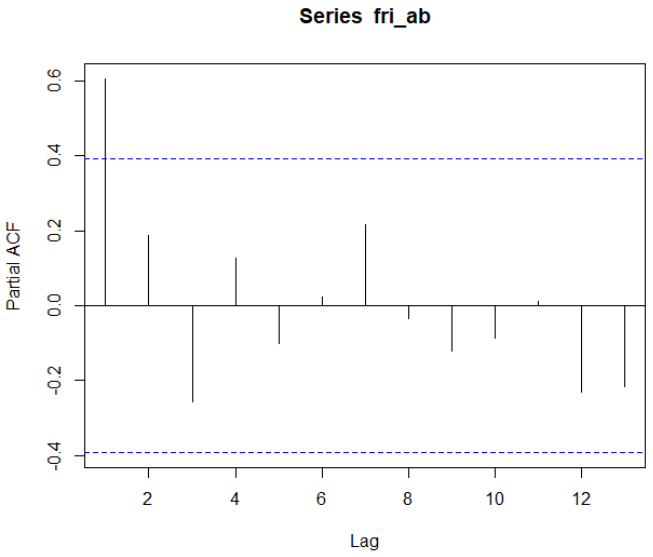
- ARIMA 모형 적합

- Likes(우호도)

1) Adf검정 및 acf,pacf

```
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
     lag    ADF p.value
[1,]   0 -0.816   0.377
[2,]   1 -0.133   0.597
[3,]   2 -0.395   0.521
Type 2: with drift no trend
     lag   ADF p.value
[1,]   0 -2.28   0.231
[2,]   1 -1.59   0.478
[3,]   2 -2.21   0.255
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -3.50  0.0634
[2,]   1 -2.35  0.4221
[3,]   2 -3.19  0.1185
----
```





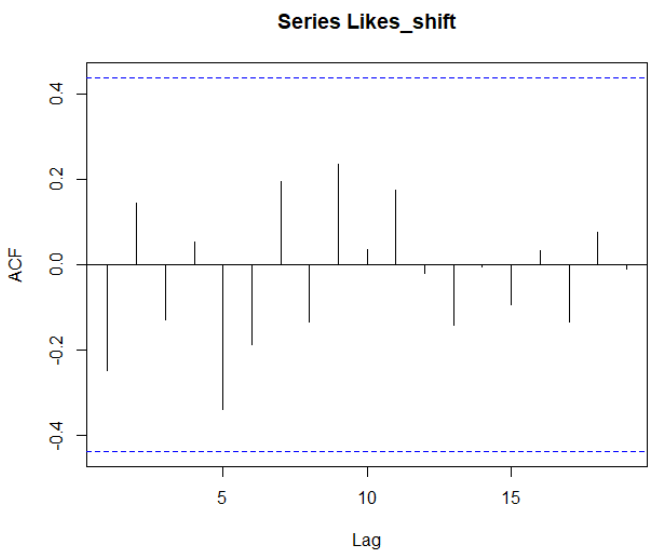
모형 테스트

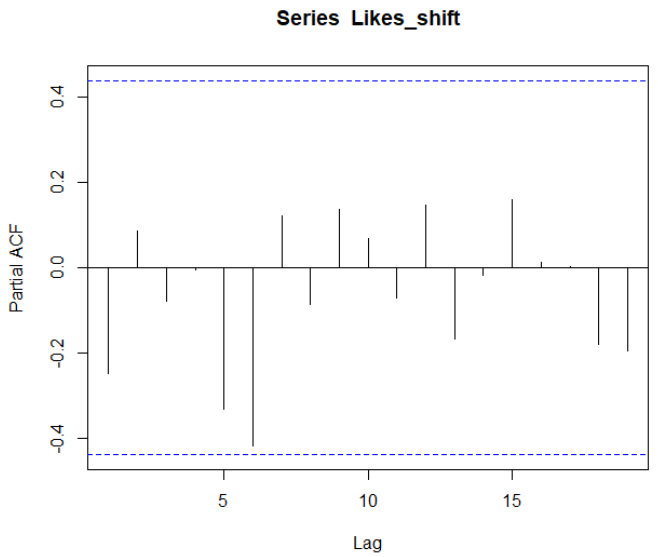
- ARIMA 모형 적합

- Likes(우호도)

2) Adf검정 및 acf,pacf(차분 후)

```
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
     lag   ADF p.value
[1,]   0 -5.22  0.0100
[2,]   1 -2.10  0.0375
[3,]   2 -1.80  0.0720
Type 2: with drift no trend
     lag   ADF p.value
[1,]   0 -5.30   0.010
[2,]   1 -2.28   0.231
[3,]   2 -2.02   0.325
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -5.71   0.010
[2,]   1 -2.53   0.357
[3,]   2 -2.23   0.467
```



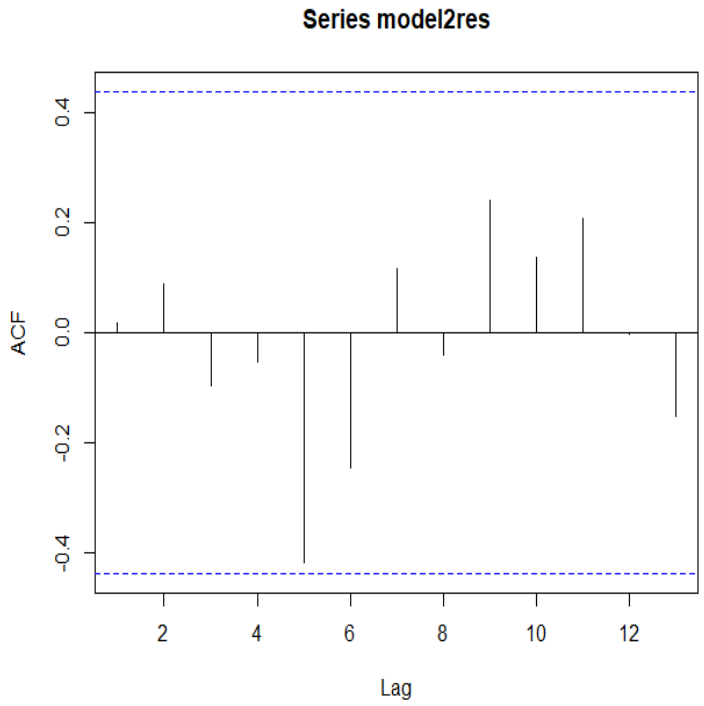


모형 테스트

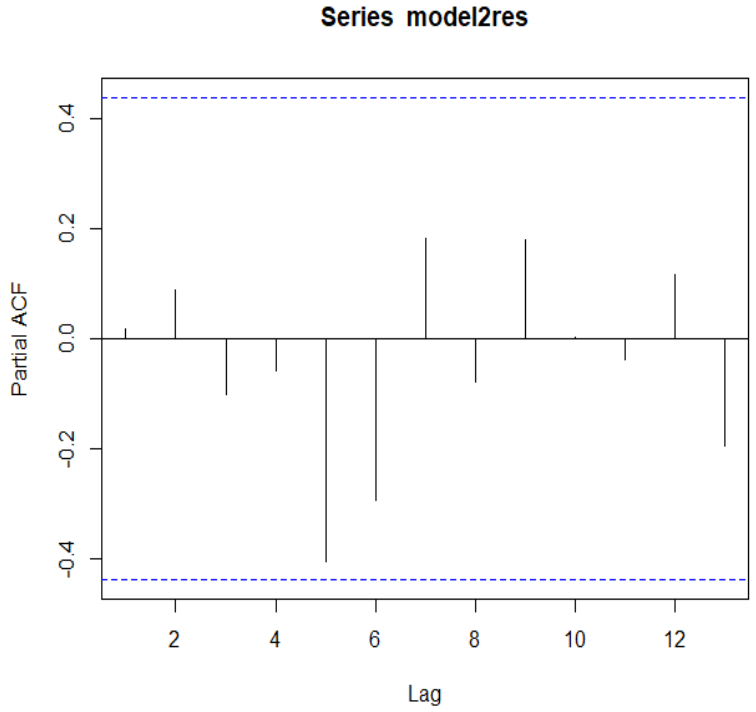
- ARIMA 모형 적합

- Likes(우호도)

3) 임의로 ARIMA(1,1,0)을 적용한다.




```
        Box-Ljung test

data:  model2res
X-squared = 0.0063529, df = 1, p-value = 0.9365
```



모형 테스트

- Summary

- 방한관광객수 & 환율 단독모델

	26차시	27차시	28차시	29차시	30차시
참값	3518792	2747750	2280434	1837782	2297893
예측값	3473058	2783635	2473223	1816079	2269820
오차율	1.2%	1.3%	8.45%	1.18%	1.22%

MAPE = 2.67

- 방한관광객수 & 환율 & 우호도 복합모델

	26차시	27차시	28차시	29차시	30차시
참값	3518792	2747750	2280434	1837782	2297893
예측값	3404289	2780888	2471284	1814679	2268027
오차율	3.2%	1.2%	8.36%	1.25%	1.29%

MAPE = 3.06

결론

- 방한관광객수는 환율, 우호도와 모두 교차상관성이 존재

- 방한관광객수 – 환율 단독모델이
  방한관광객수 – 환율&우호도 복합모델보다
  예측력이 우수

- Log 변환을 하여 모형을 적합한 결과 더 우수한 결과가 나옴

감사합니다