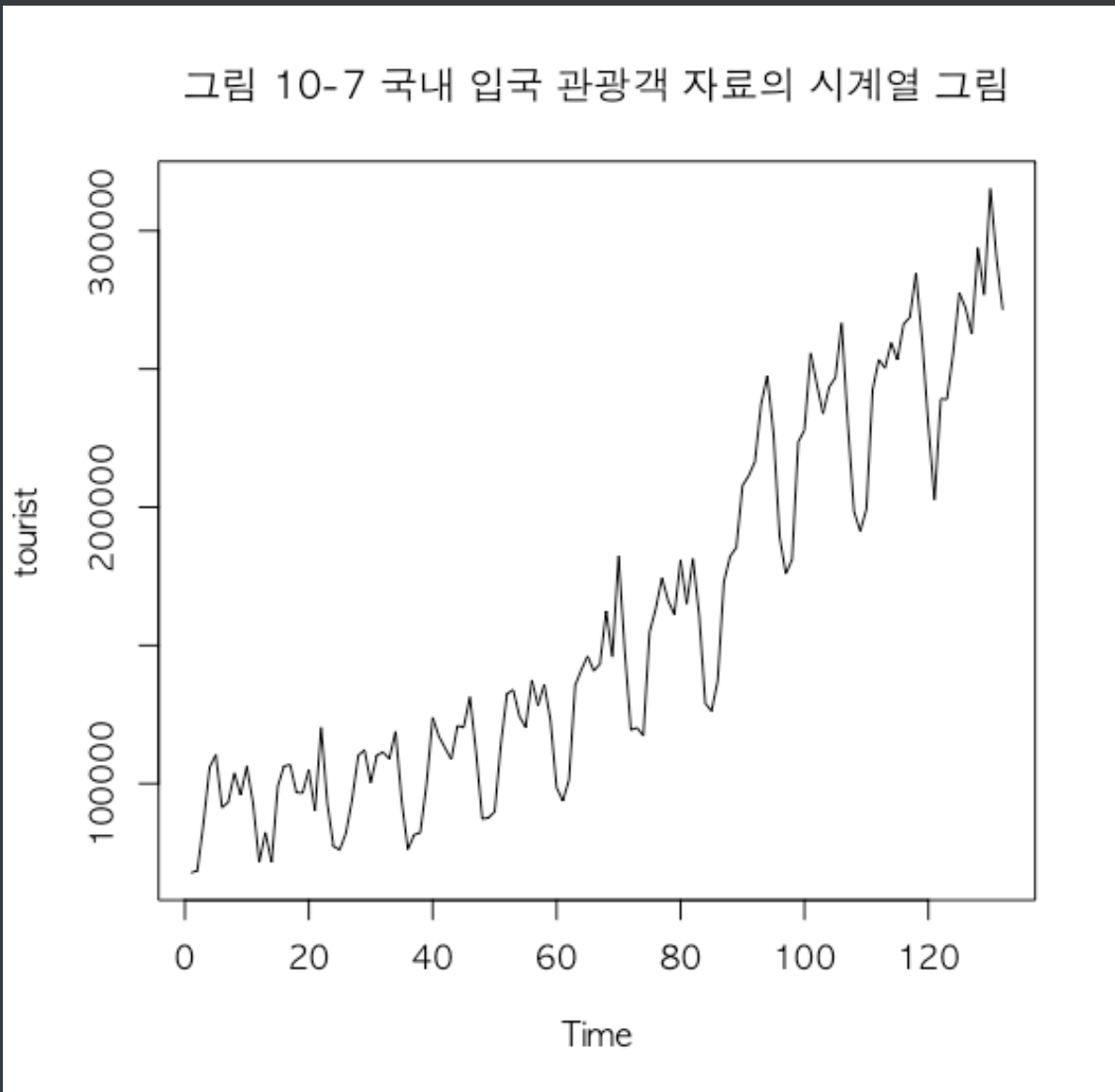


HW 10

2021234640 이종현

Figure 10-7

이번 챕터에서는 ARIMA에 계절 요소를 적합할 수 있는 seasonal 요소를 추가하는 방법을 배웠다. 아래는 국내 입국 관광객 자료이다.



계절성, 추세 성분이 눈에 띈다. 또한 분산이 시간에 따라 커진다.

Figure 10-8

로그 변환을 취하여 분산을 안정화하자.

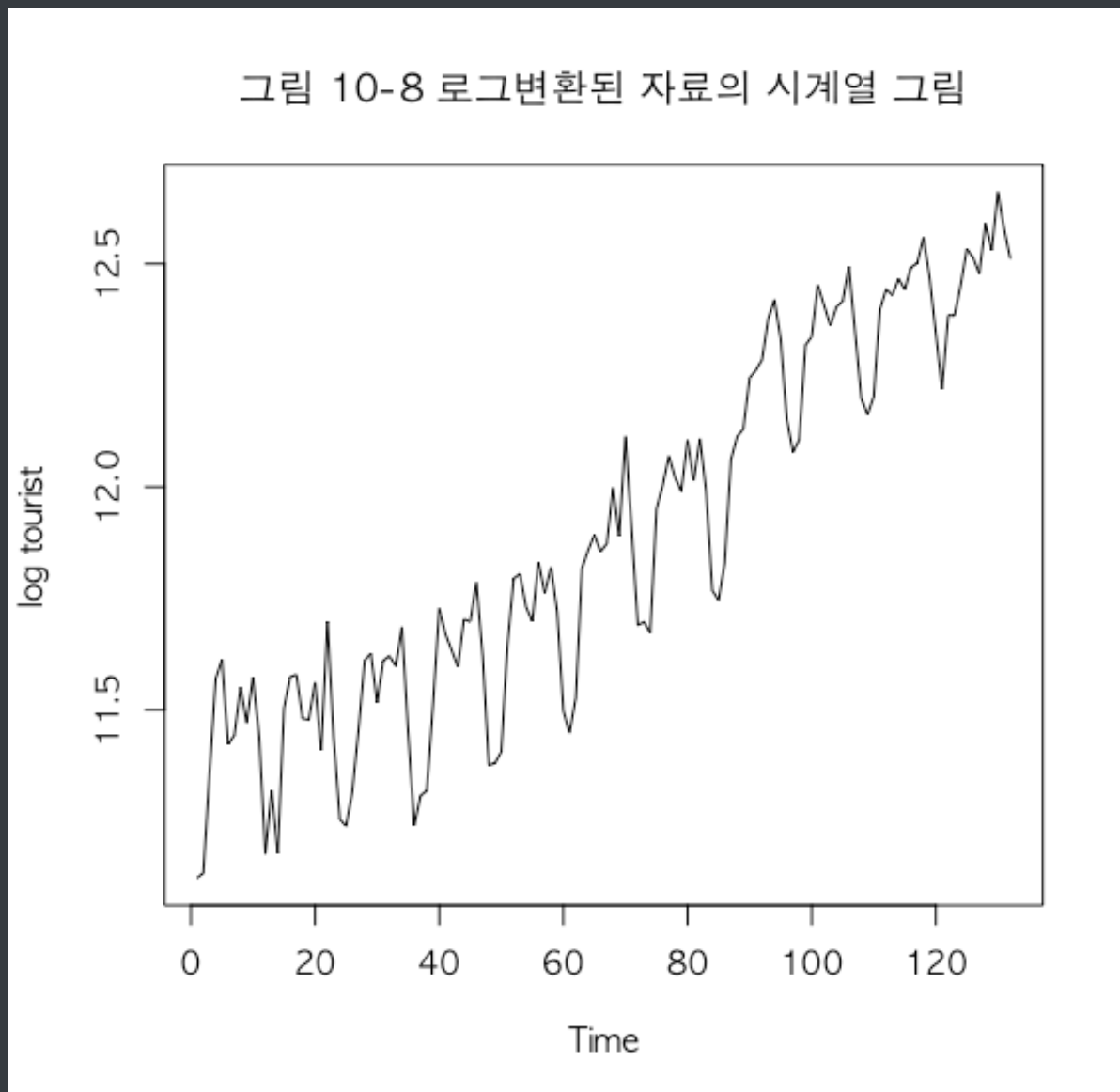
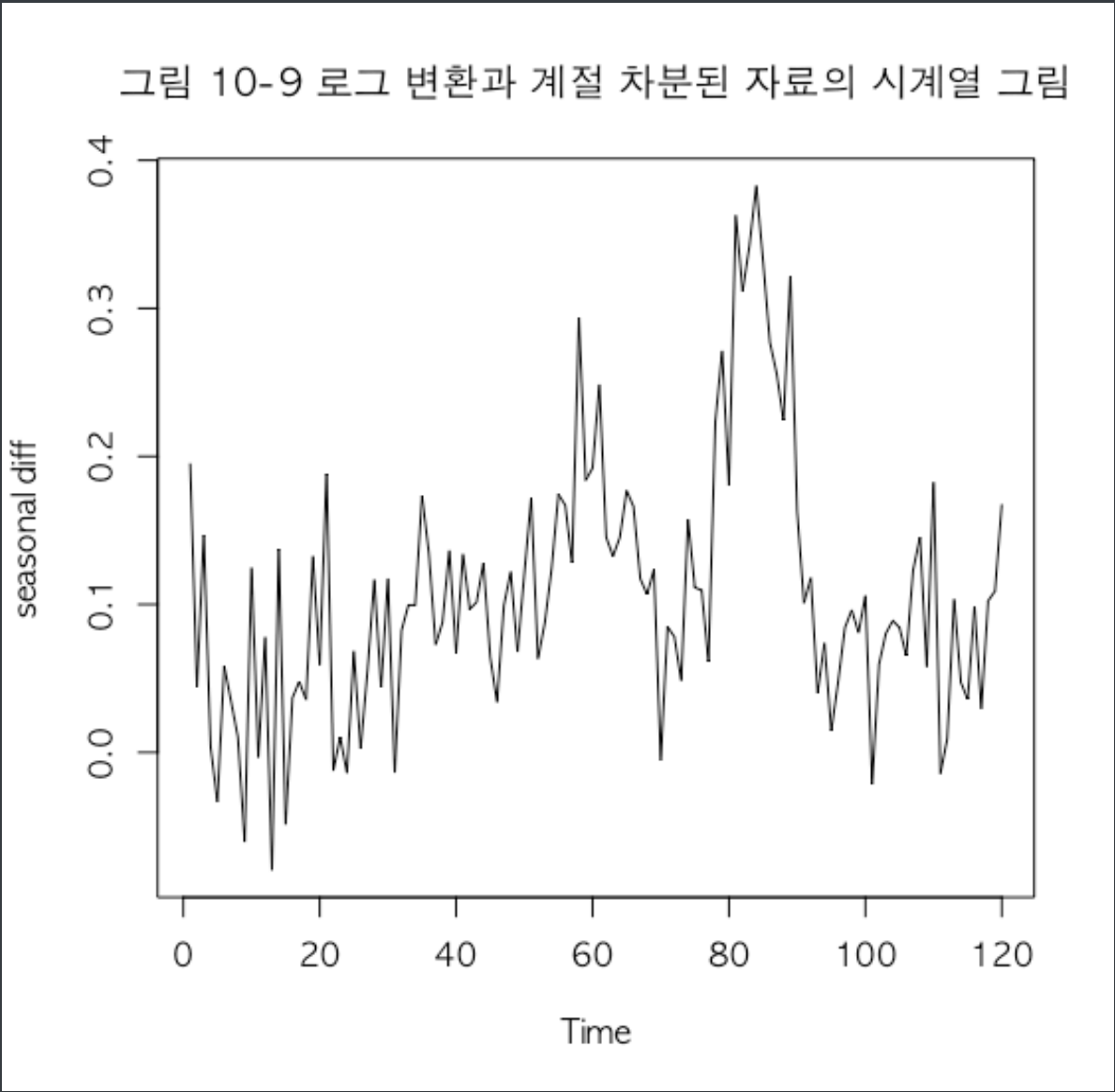
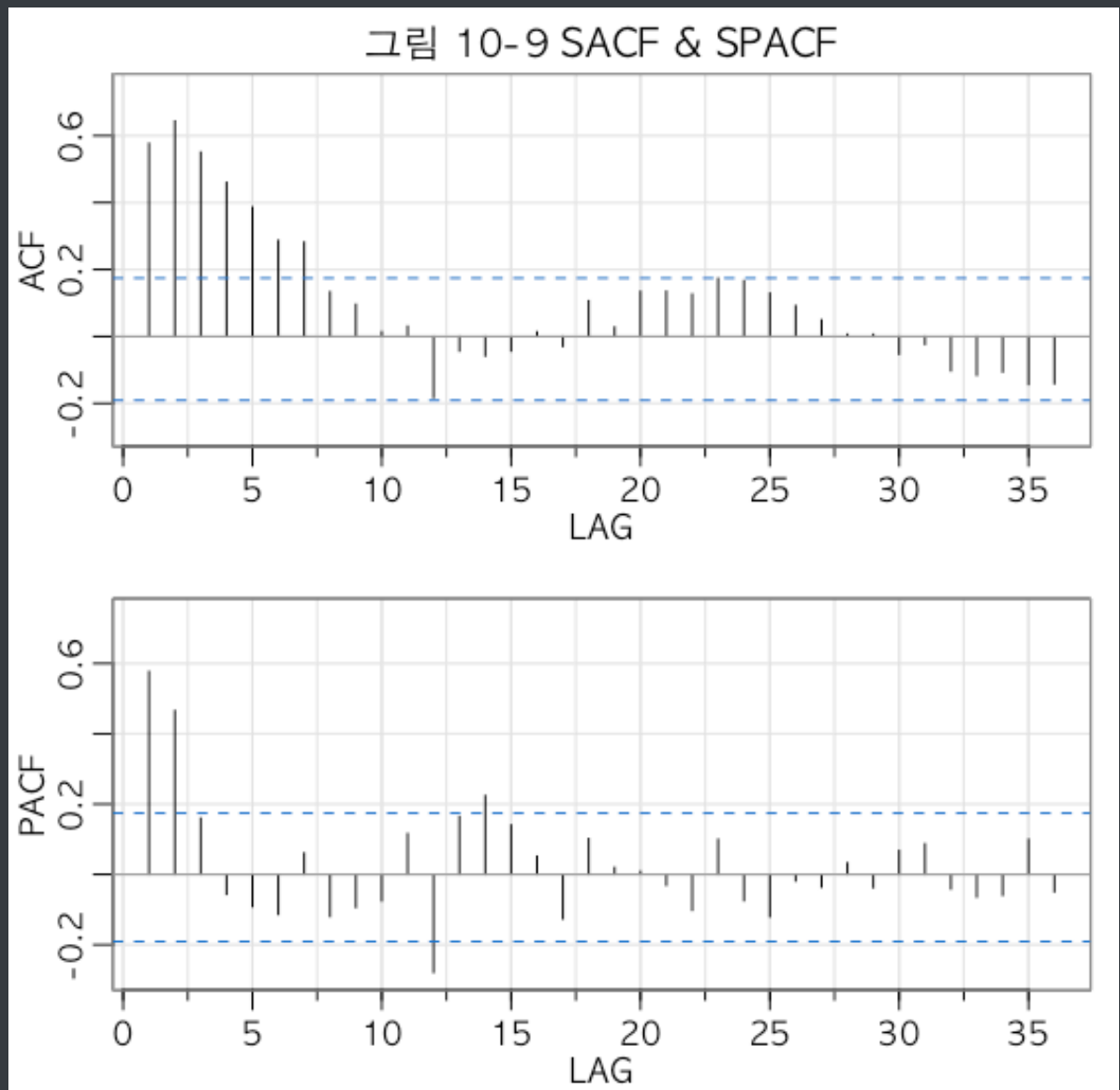


Figure 10-9

먼저 계절 차분을 진행해보자. lag 12에 대해 차분을 진행한다.

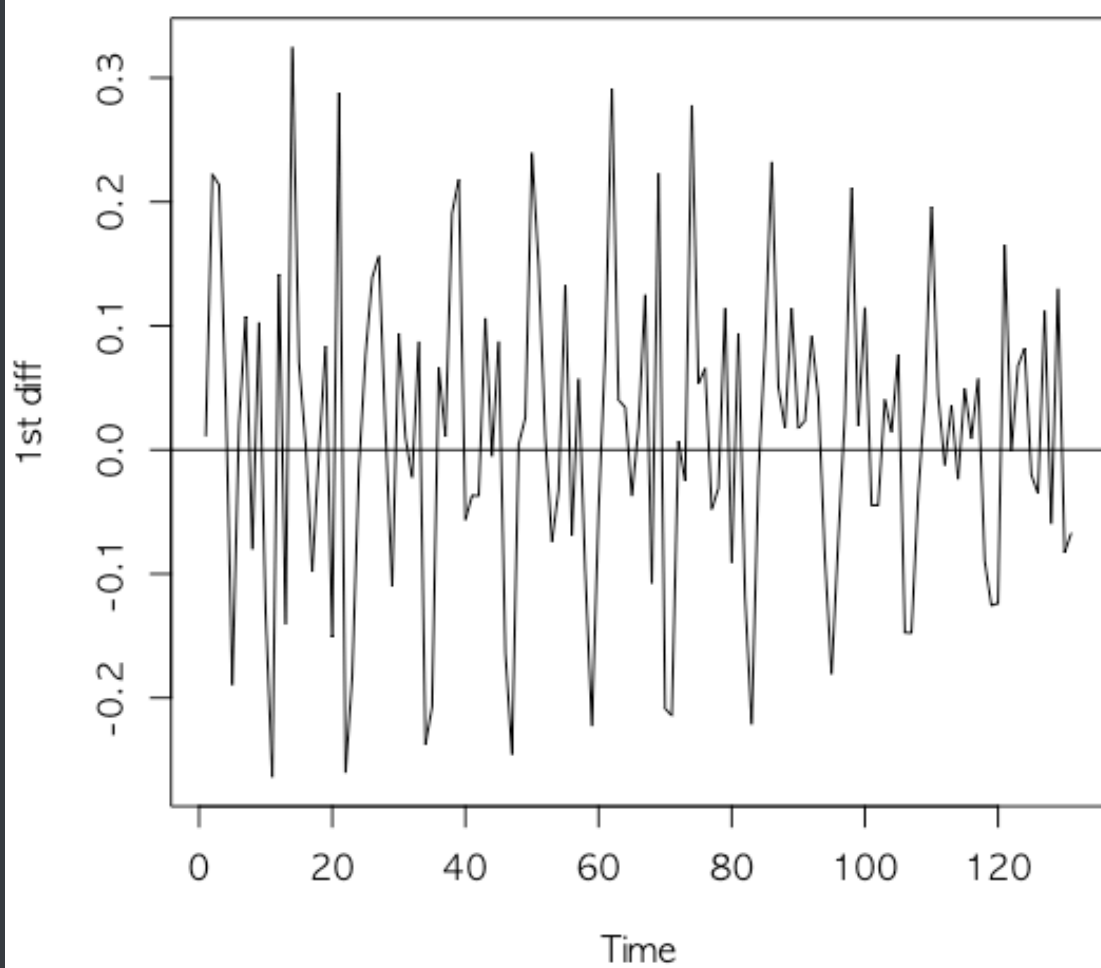


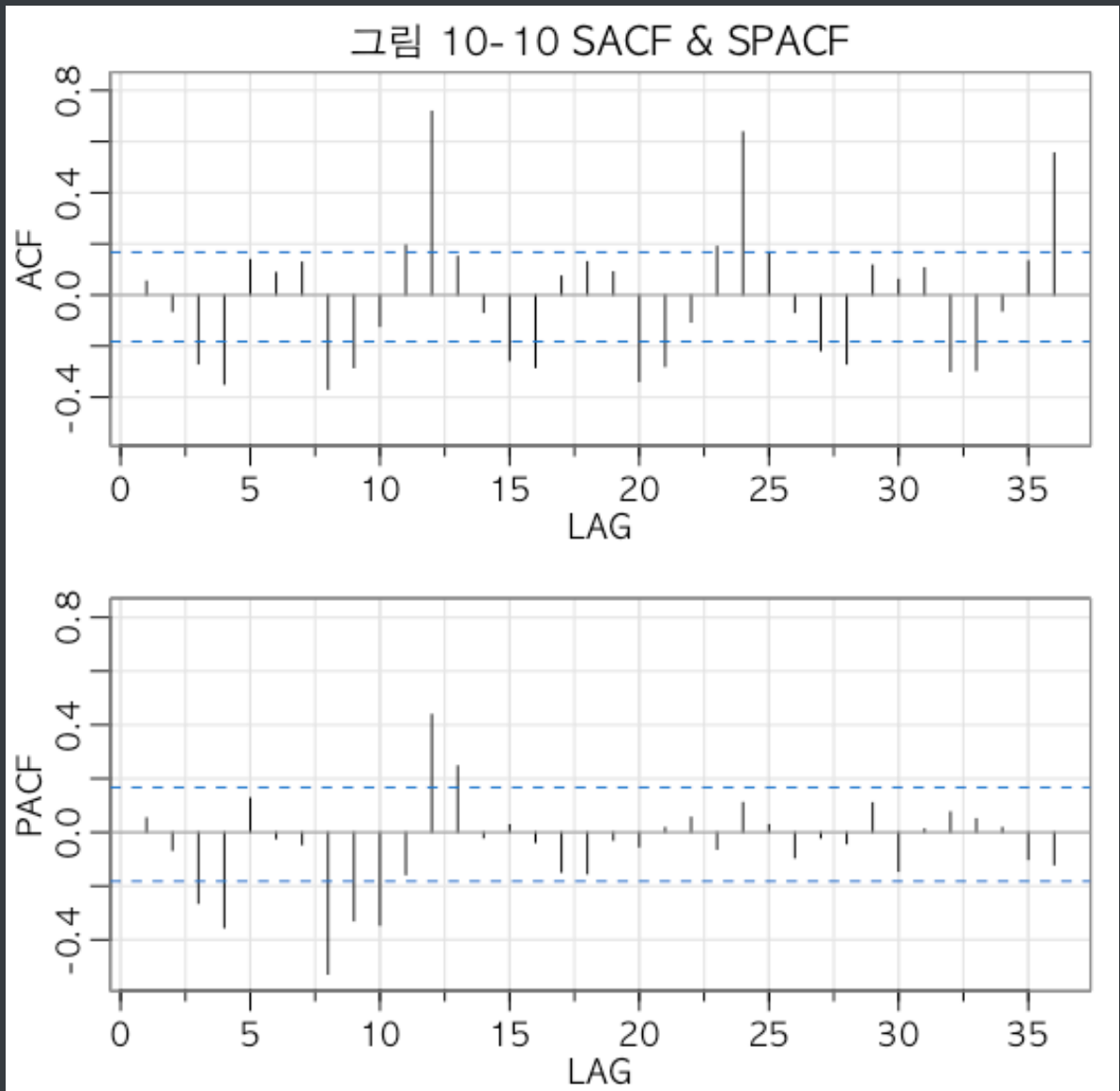


아직 정상 시계열은 아니다. 그러나 계절성은 많이 줄어든 것을 확인할 수 있다. ACF가 천천히 줄어들고 PACF가 2에서 절단되는 모습을 볼 수 있다. AR 1, 2를 고려해볼만하다.

Figure 10-10

그림 10-10 로그 변환과 1차 차분된 자료의 시계열 그림



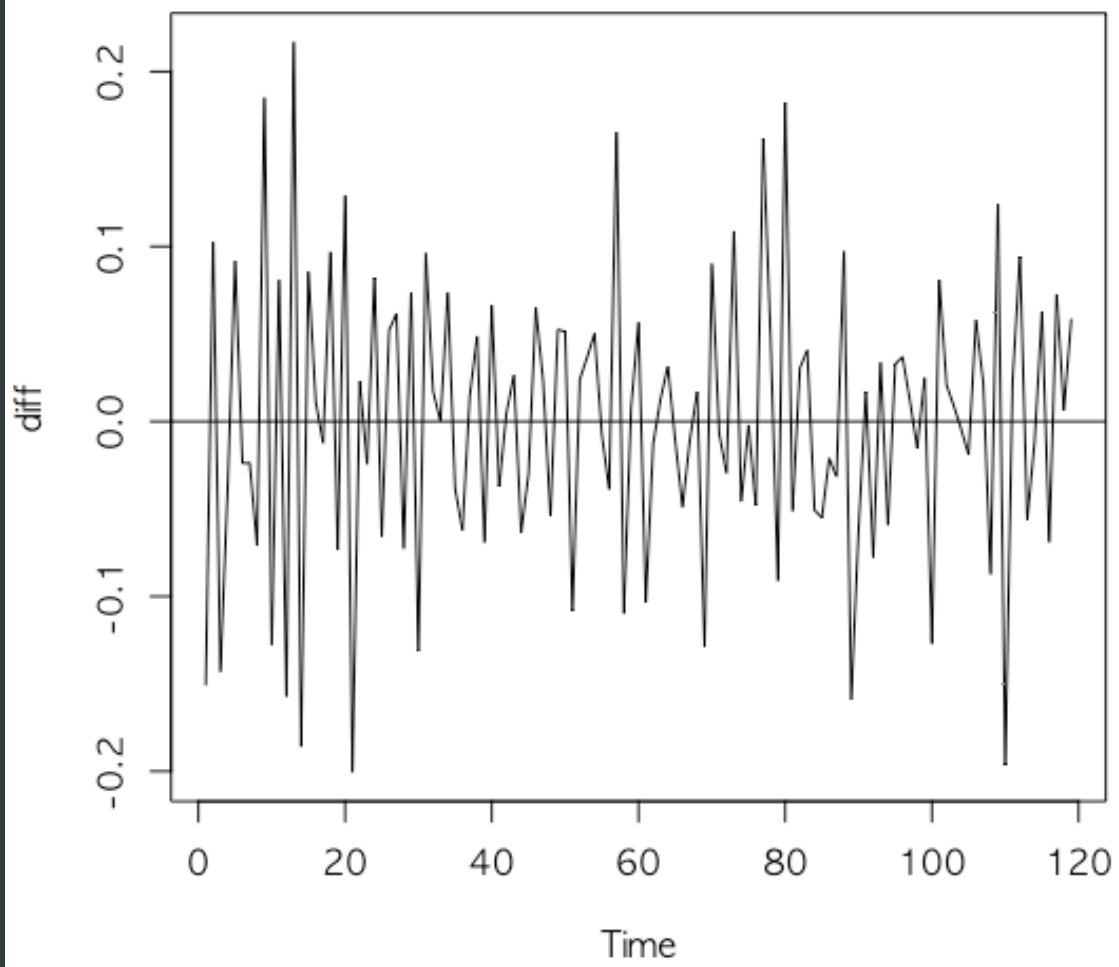


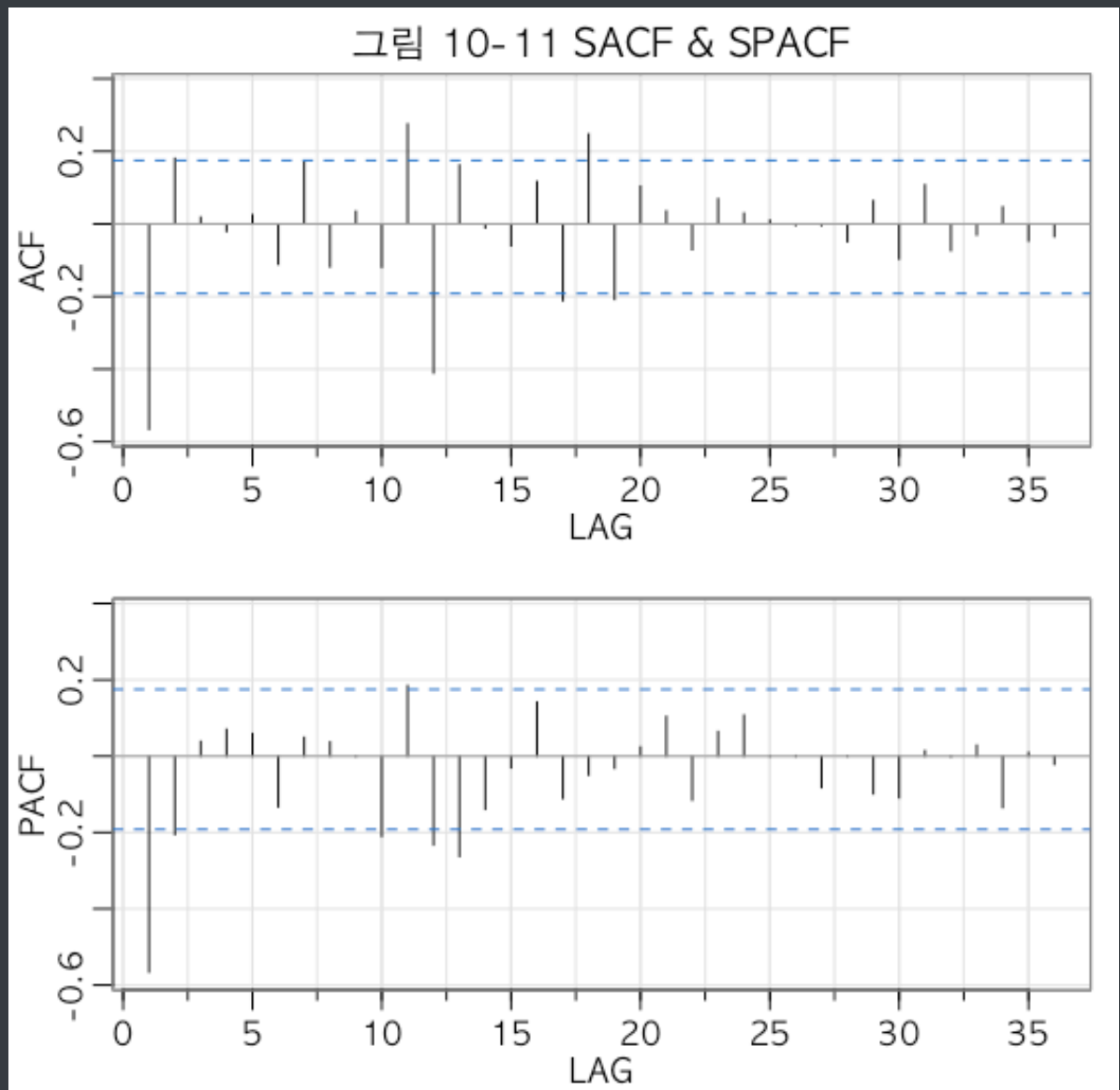
이번에는 1차 차분을 진행하였다. ACF의 12, 24, 36 에서 계절성의 특징 패턴이 보인다.

Figure 10-11

계절 차분과 1차 차분을 동시에 적용해보자.

그림 10-11 로그 변환과 계절 및 1차 차분된 자료의 시계열 그



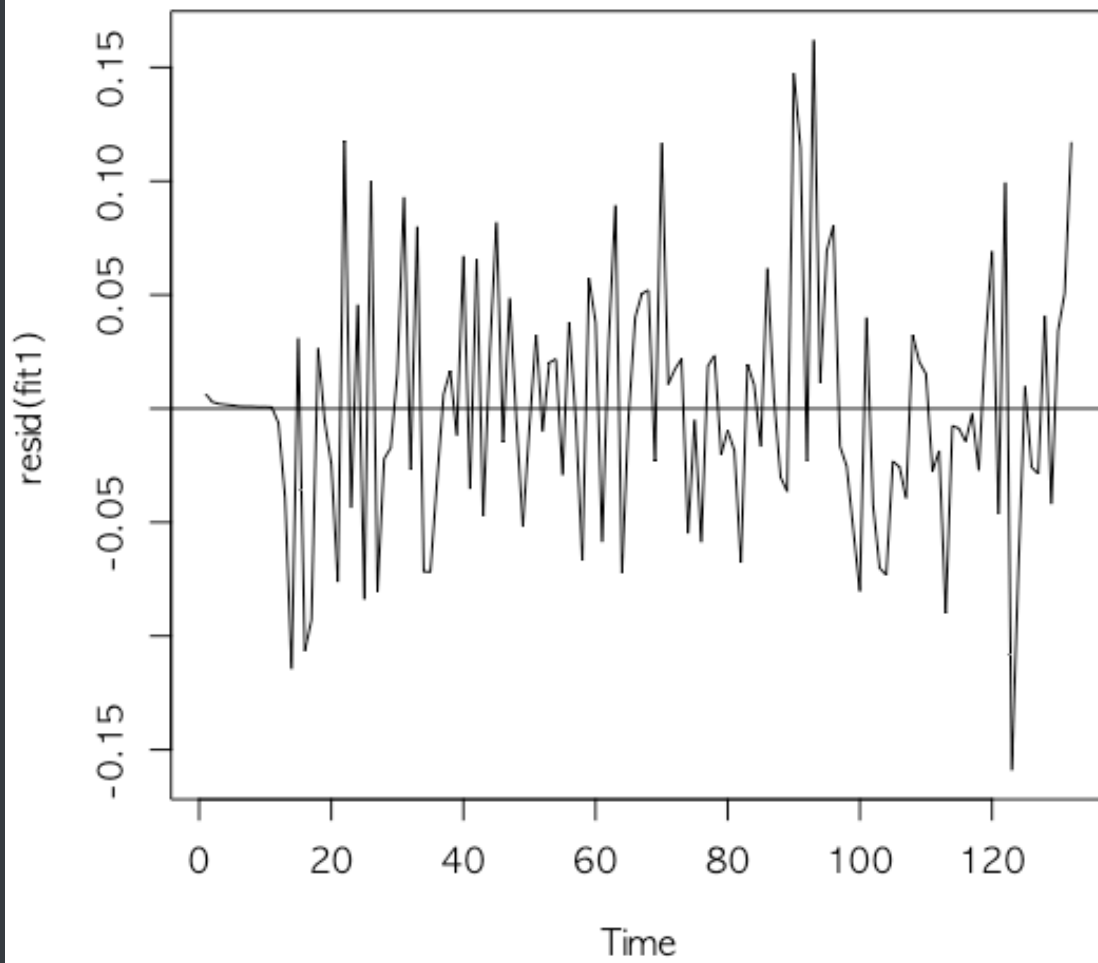


이제까지 결과 중 가장 정상 시계열에 가깝게 되었다. 이 결과를 바탕으로 ARIMA를 적합해보자.

Figure 10-12

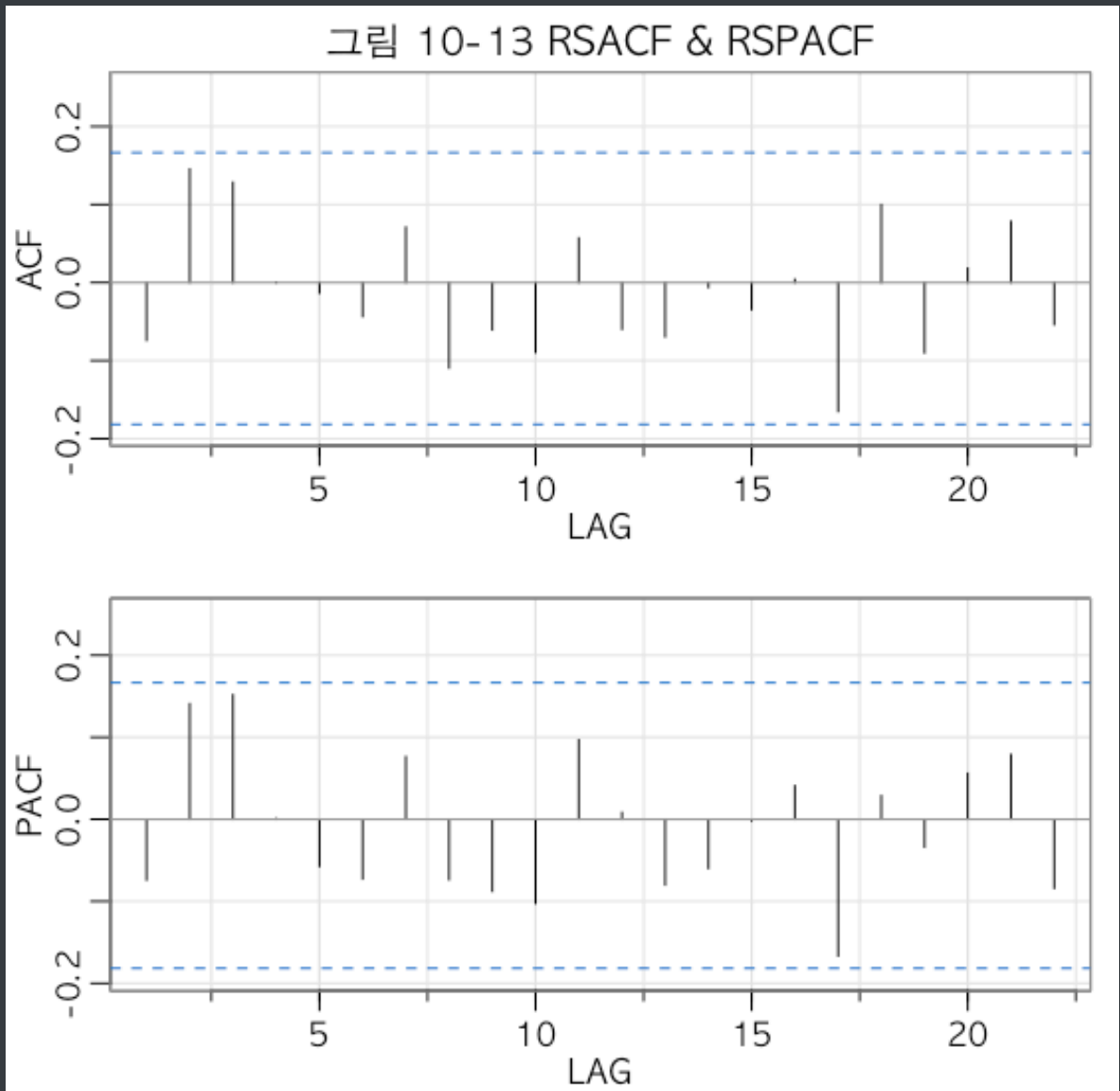
ARIMA(0, 1, 1), (0, 1, 1) 모델을 적합하였다.

그림 10-12 잔차의 시계열 그림



이때의 잔차는 다음과 같다.

Figure 10-13

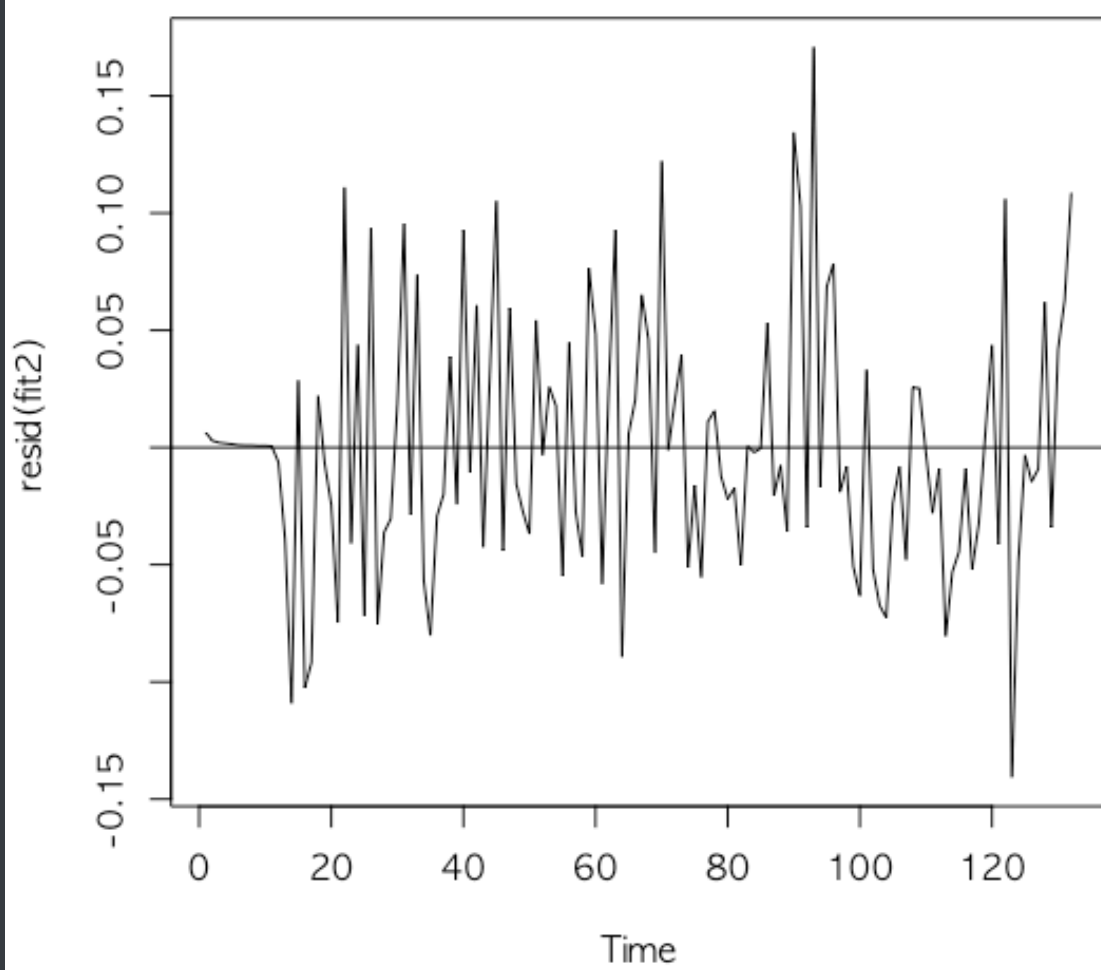


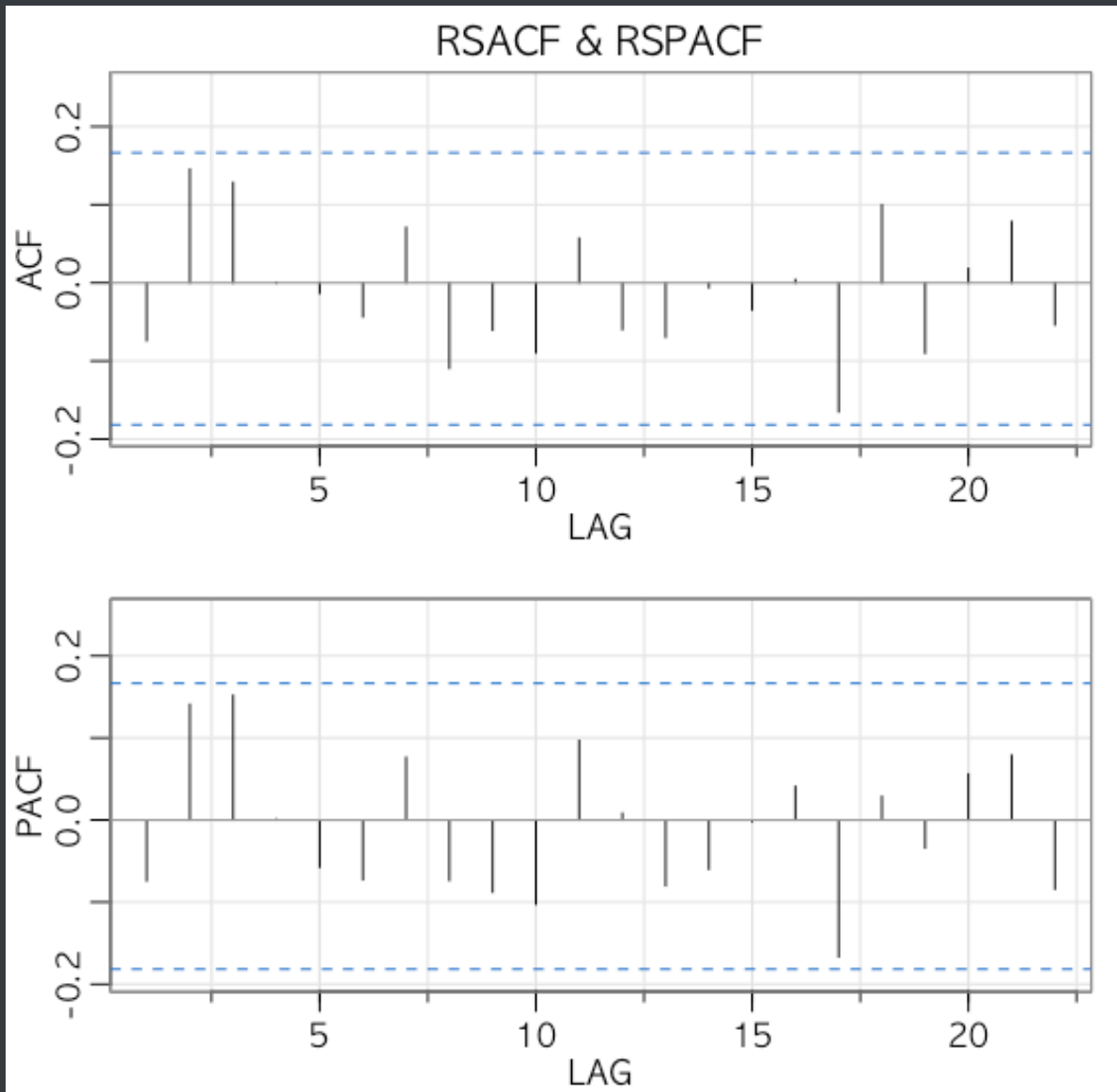
잔차가 상당히 안정적인 패턴을 보이고 있다. 모델링이 잘 되었다고 판단할 수 있겠다. 교재에서는 $ARIMA(0, 1, 1)$, $(1, 1, 0)$ 을 고려할 수 있다고 언급하고 있다.

Figure 10–14

$ARIMA(0, 1, 1)$, $(1, 1, 0)$ 의 적합 결과는 다음과 같다.

잔차의 시계열 그림





이 역시 잔차가 안정적이기 때문에 모형으로 선택할 여지가 있다고 보인다.

Appendix: R code

```
rm(list=ls())

setwd("~/Workspace/2022-Fall_TimeSeriesAnalysis/data/")
par(family="AppleGothic")

# 10.5
library(astsa)
tour = scan("tourist.txt")
```

```

ts.plot(tour, ylab="tourist", main="그림 10-7 국내 입국 관광객 자료의 시계열 그림")

ltour = log(tour)
ts.plot(ltour, ylab="log tourist", main="그림 10-8 로그변환된 자료의 시계열 그림")

d12tour = diff(ltour, lag=12)
ts.plot(d12tour, ylab="seasonal diff", main="그림 10-9 로그 변환과 계절 차분된 자료의 시계열 그림")
acf2(d12tour, max.lag=36, main="그림 10-9 SACF & SPACF")

d1tour = diff(ltour, lag=1)
ts.plot(d1tour, ylab="1st diff", main="그림 10-10 로그 변환과 1차 차분된 자료의 시계열 그림")
abline(h=0)
acf2(d1tour, max.lag=36, main="그림 10-10 SACF & SPACF")

d1_12tour = diff(d12tour)
ts.plot(d1_12tour, ylab="diff", main="그림 10-11 로그 변환과 계절 및 1차 차분된 자료의 시계열 그림")
abline(h=0)
acf2(d1_12tour, max.lag=36, main="그림 10-11 SACF & SPACF")

fit1 = arima(ltour, order=c(0, 1, 1), seasonal=list(order=c(0, 1, 1),
period=12))
fit1

ts.plot(resid(fit1), main="그림 10-12 잔차의 시계열 그림")
abline(h=0)

acf2(resid(fit1), main="그림 10-13 RSACF & RSPACF")
Box.test(resid(fit1), lag=6, type="Ljung", fitdf=0)

fit2 = arima(ltour, order=c(0,1,1), seasonal=list(order=c(1,1,0),
period=12))

```

```
fit2
```

```
ts.plot(resid(fit2), main="잔차의 시계열 그림")
```

```
abline(h=0)
```

```
acf2(resid(fit1), main="RSACF & RSPACF")
```

Appendix: Python code

```
In [1]: import math
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
plt.rcParams['font', family='AppleGothic']
plt.rcParams['axes.unicode_minus'] = False

from statsmodels.tsa.arima_model import ARIMA
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
```

```
In [2]: # Example 9.1
z = []

with open('../data/tourist.txt') as f:
    for line in f.readlines():
        for elem in line.rstrip().split(" "):
            if len(elem):
                z.append(float(elem))

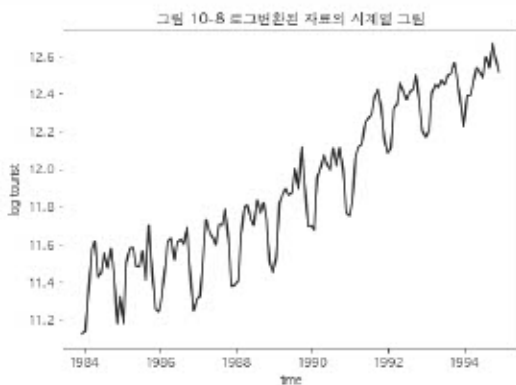
index = pd.date_range(start="1984", periods=len(z), freq="MS")
data = pd.Series(z, index)

fig, ax = plt.subplots(figsize=(7, 5))
ax.plot(data, 'black')
ax.set_xlabel("time")
ax.set_ylabel("tourist")
ax.set_title("그림 10-7 국내 입국 관광객 자료의 시계열 그림")
plt.show()
```



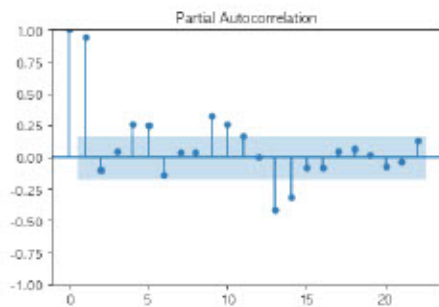
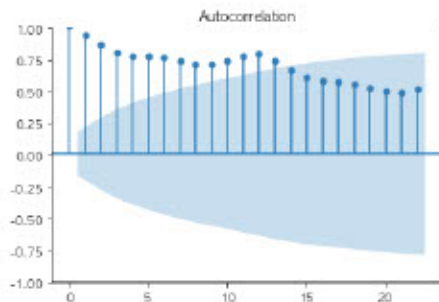
```
In [3]: ltour = np.log(data)

fig, ax = plt.subplots(figsize=(7, 5))
ax.plot(ltour, 'black')
ax.set_xlabel("time")
ax.set_ylabel("log tourist")
ax.set_title("그림 10-8 로그변환된 자료의 시계열 그림")
plt.show()
```



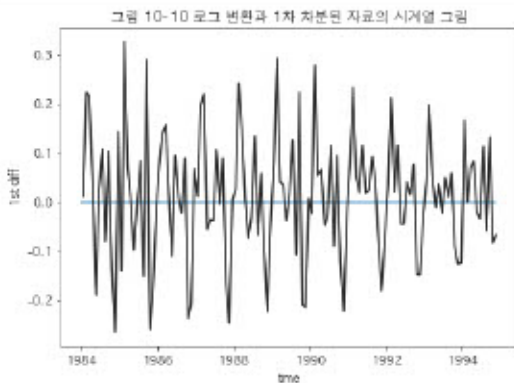
```
In [4]: plot_acf(ltour)
plot_pacf(ltour)
plt.show()
```

/Users/jonghyun/.local/lib/python3.9/site-packages/statsmodels/graphics/tseplots.py:348: FutureWarning: The default method 'yw' can produce PACF values outside of the [-1,1] interval. After 0.13, the default will change to unadjusted Yule-Walker ('yw'). You can use this method now by setting method='yw'.
warnings.warn(



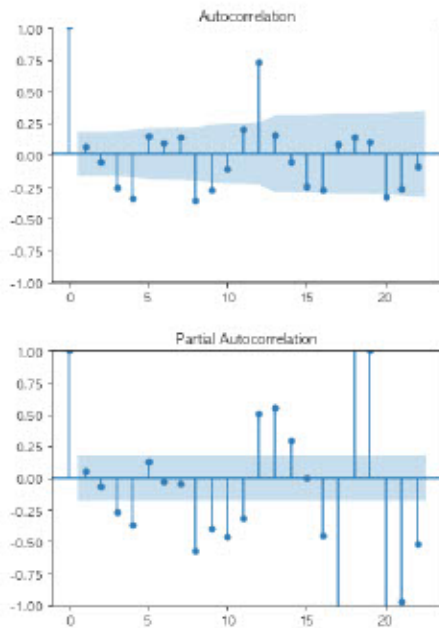
```
In [5]: dltour = ltour.diff(1)

fig, ax = plt.subplots(figsize=(7, 5))
ax.plot(dltour, 'black')
ax.set_xlabel("time")
ax.set_ylabel("1st diff")
ax.set_title("그림 10-10 로그 변환과 1차 차분된 자료의 시계열 그림")
ax.hlines(0, dltour.index.min(), dltour.index.max())
plt.show()
```



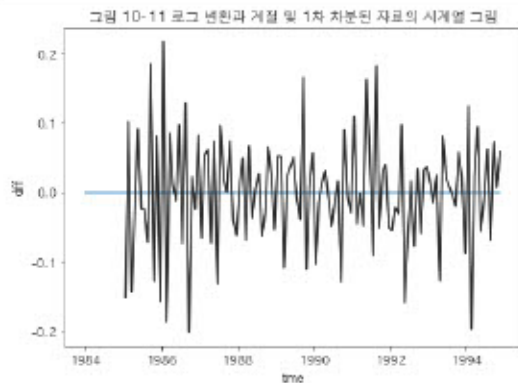

```
In [6]: plot_acf(dltour[1:])
plot_pacf(dltour[1:])
plt.show()
```

/Users/jonghyun/.local/lib/python3.9/site-packages/statsmodels/graphics/tseplots.py:348: FutureWarning: The default method 'yw' can produce PACF values outside of the [-1,1] interval. After 0.13, the default will change to unadjusted Yule-Walker ('yw'). You can use this method now by setting method='yw'.



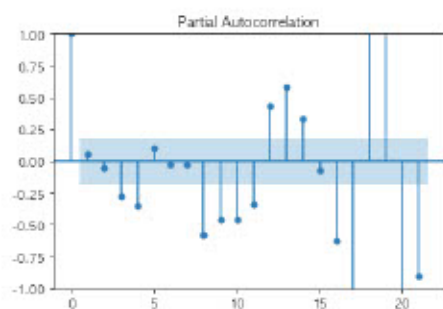
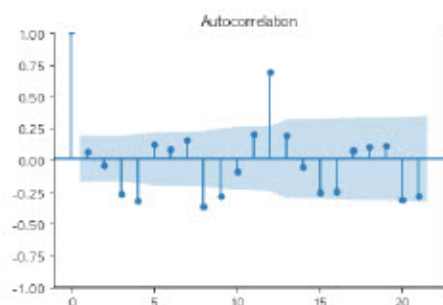
```
In [7]: dl_l2tour = dltour.diff(12)

fig, ax = plt.subplots(figsize=(7, 5))
ax.plot(dl_l2tour, 'black')
ax.set_xlabel("time")
ax.set_ylabel("diff")
ax.set_title("그림 10-11 로그 변환과 계절 및 1차 차분된 자료의 시계열 그림")
ax.hlines(0, dl_l2tour.index.min(), dl_l2tour.index.max())
plt.show()
```



```
In [8]: plot_acf(dltour[12:])
plot_pacf(dltour[12:])
plt.show()
```

/Users/jonghyun/.local/lib/python3.9/site-packages/statsmodels/graphics/tseplots.py:348: FutureWarning: The default method 'yw' can produce PACF values outside of the $[-1,1]$ interval. After 0.13, the default will change to unadjusted Yule-Walker ('yw_m'). You can use this method now by setting method='yw_m'.
warnings.warn(



```
In [18]: from statsmodels.tsa.statespace.sarimax import SARIMAX

fit1 = SARIMAX(data, order=(0, 1, 1), seasonal_order=(0, 1, 1, 12)).fit()
resid = fit1.resid

plt.plot(resid)
plt.title("그림 10-12 잔차의 시계열 그림")
plt.hlines(0, resid.index.min(), resid.index.max(), color="black")
plt.show()
```

RUNNING THE L-BFGS-B CODE

```

* * *

Machine precision = 2.220D-16
N =          3      M =          10

At X0          0 variables are exactly at the bounds
At iterate   0   f= 9.66365D+00   |proj g|= 2.66652D-01
At iterate   5   f= 9.61560D+00   |proj g|= 1.06702D-05

* * *

Tit  = total number of iterations
Tnf  = total number of function evaluations
Tnint = total number of segments explored during Cauchy searches
Skip  = number of BFGS updates skipped
Nact  = number of active bounds at final generalized Cauchy point
Projg = norm of the final projected gradient
F     = final function value

```

```

* * *

      N      Tit      Tnf      Tnint      Skip      Nact      Projg      F
      3         6         8         1         0         0   1.067D-05   9.616D+00
F = 9.6155982262041153

```

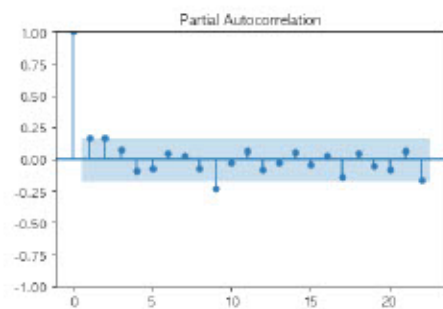
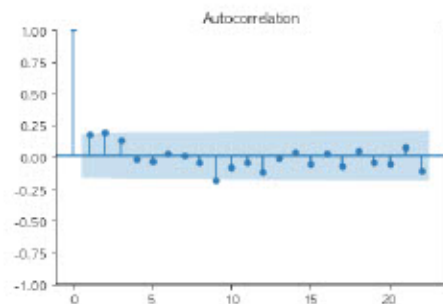
CONVERGENCE: REL_REDUCTION_OF_F_<=_FACTR*EPSMCH

This problem is unconstrained.



```
In [19]: plot_acf(resid)
plot_pacf(resid)
plt.show()
```

/Users/jonghyun/.local/lib/python3.9/site-packages/statsmodels/graphics/tseplots.py:348: FutureWarning: The default method 'yw' can produce PACF values outside of the $[-1,1]$ interval. After 0.13, the default will change to unadjusted Yule-Walker ('yw'). You can use this method now by setting method='yw'.
warnings.warn(



```
In [20]: fit2 = SARIMAX(data, order=(0, 1, 1), seasonal_order=(1, 1, 0, 12)).fit()
resid = fit2.resid

plt.plot(resid)
plt.title("그림 10-12 잔차의 시계열 그림")
plt.hlines(0, resid.index.min(), resid.index.max(), color="black")
plt.show()
```

RUNNING THE L-BFGS-B CODE

```

* * *

Machine precision = 2.220D-16
N =          3      M =          10

At X0          0 variables are exactly at the bounds

At iterate   0    f= 9.68952D+00    |proj g|= 2.30207D-01
At iterate   5    f= 9.62915D+00    |proj g|= 7.88347D-06

* * *

Tit  = total number of iterations
Tnf  = total number of function evaluations
Tnint = total number of segments explored during Cauchy searches
Skip  = number of BFGS updates skipped
Nact  = number of active bounds at final generalized Cauchy point
Projg = norm of the final projected gradient
F     = final function value

```

```

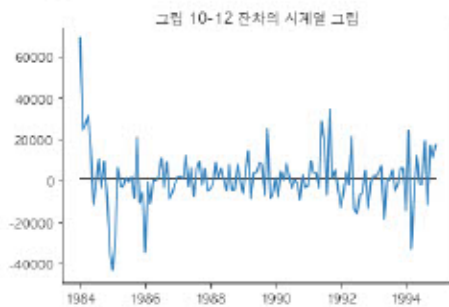
* * *

      N      Tit      Tnf  Tnint  Skip  Nact      Projg      F
      3       5       7      1      0      0    7.883D-06  9.629D+00
F = 9.6291513586809909

```

CONVERGENCE: NORM_OF_PROJECTED_GRADIENT_<=_PGTOL

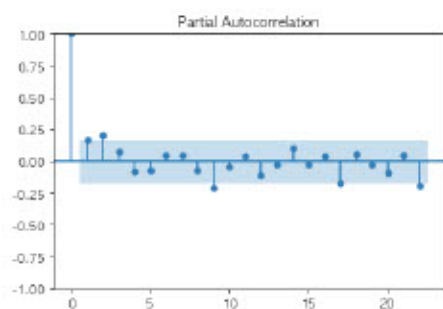
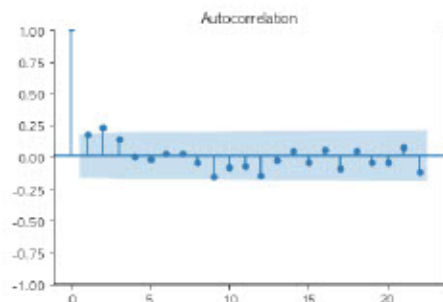
This problem is unconstrained.



```
In [21]: plot_acf(resid)
         plot_pacf(resid)
         plt.show()
```

/Users/jonghyun/.local/lib/python3.9/site-packages/statsmodels/graphics/tseplots.py:348: FutureWarning: The default method 'yw' can produce PACF values outside of the $[-1,1]$ interval. After 0.13, the default will change to unadjusted Yule-Walker ('yw'). You can use this method now by setting method='yw'.

```
warnings.warn(
```



```
In [ ]:
```