

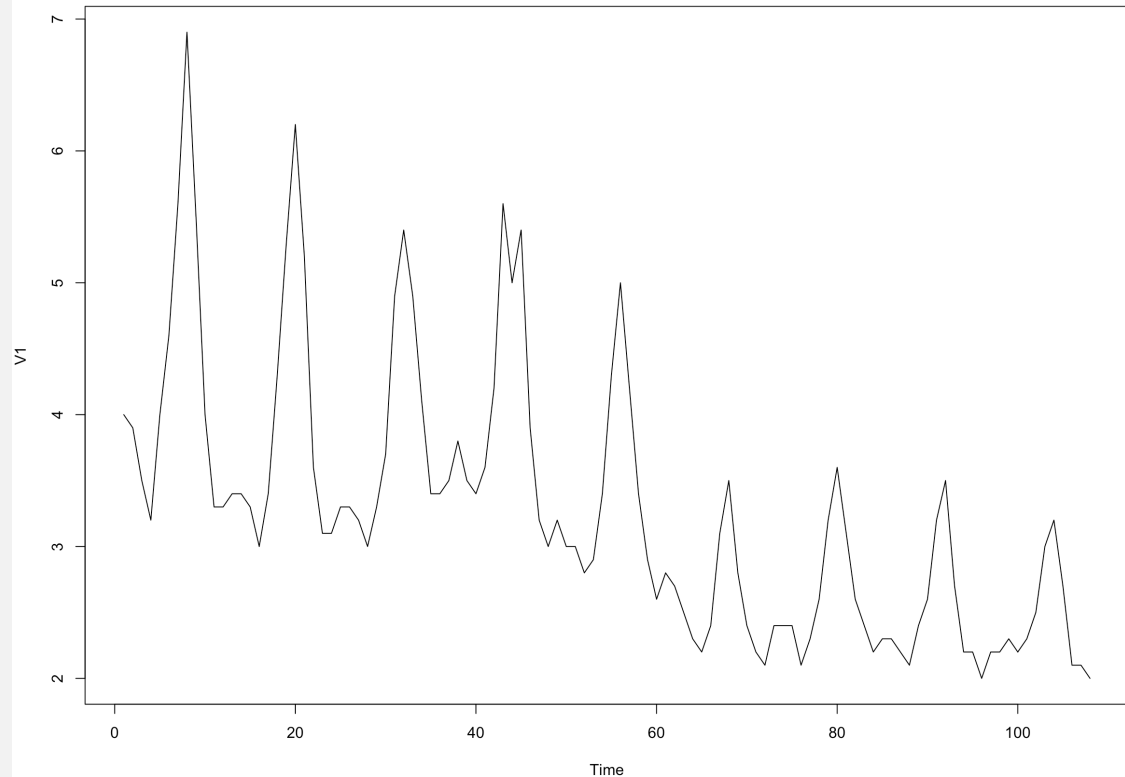
# 시계열분석 기말 포트폴리오

**2015129053 언론홍보영상학부 김형철**

## 시계열 자료의 시도표 확인

- R코드를 이용하여 시계열 시도표를 그려보니 다음과 같이 나왔다.
- 분산이 시간이 갈수록 줄어드는 듯하고, 아래로 향하는 추세도 존재하는 듯하다. 계절성도 일정부분 의심이 되는 모습이다. (참고로, 기말고사 파일은 final.csv로 이름을 변경해 사용하였다.)

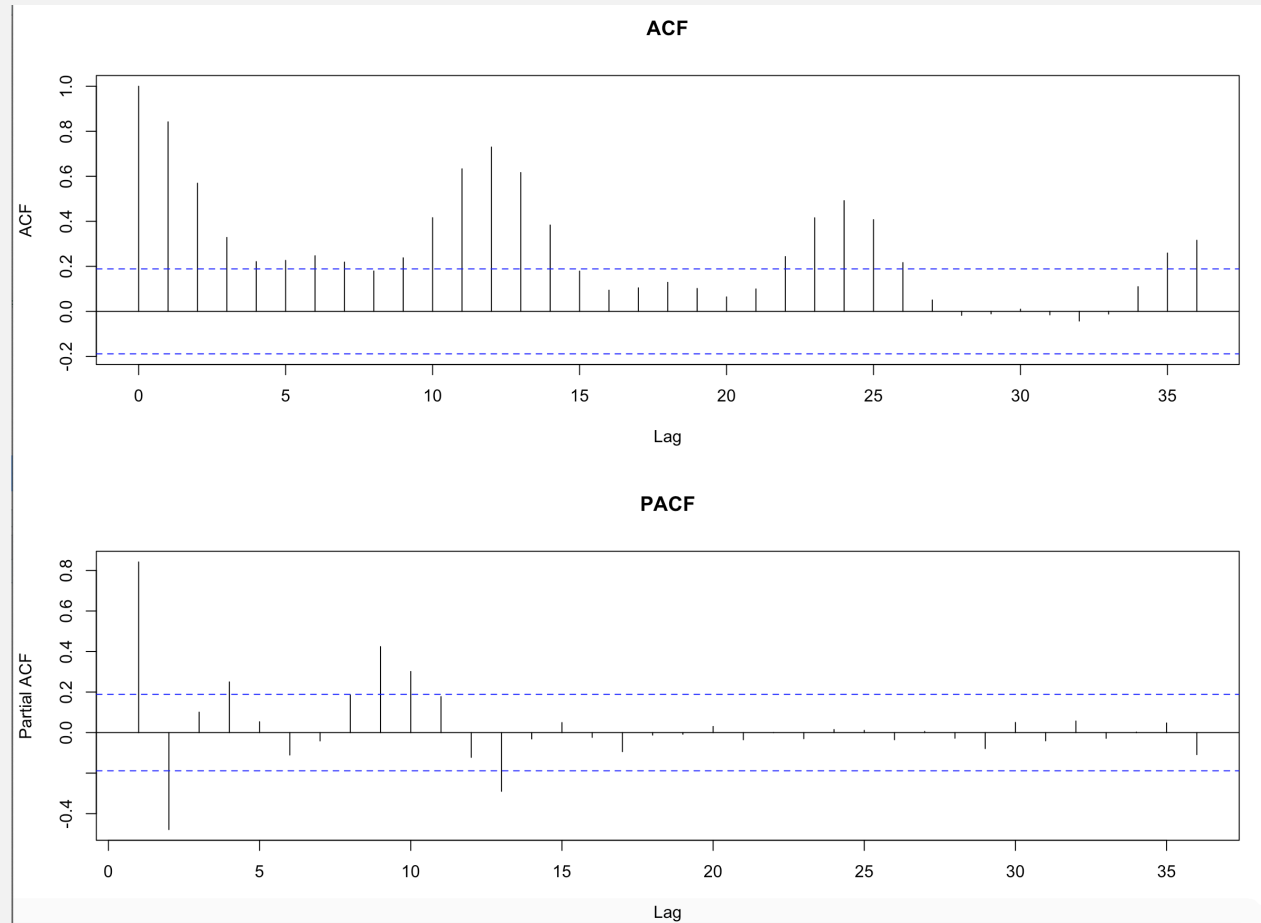
```
final<-read.csv("Desktop/final.csv", header=F)  
finalts<-ts(final)  
plot(finalts)
```



## 시계열 자료의 ACF, PACF 확인

ACF와 PACF를 확인해보면 확실히 일정부분 비정상시계열이라는 것을 확인할 수 있다.  
ACF가 지수적 감소를 하지 않고 있는 모습이 있는 등, 비정상시계열의 특징이 엿보인다.

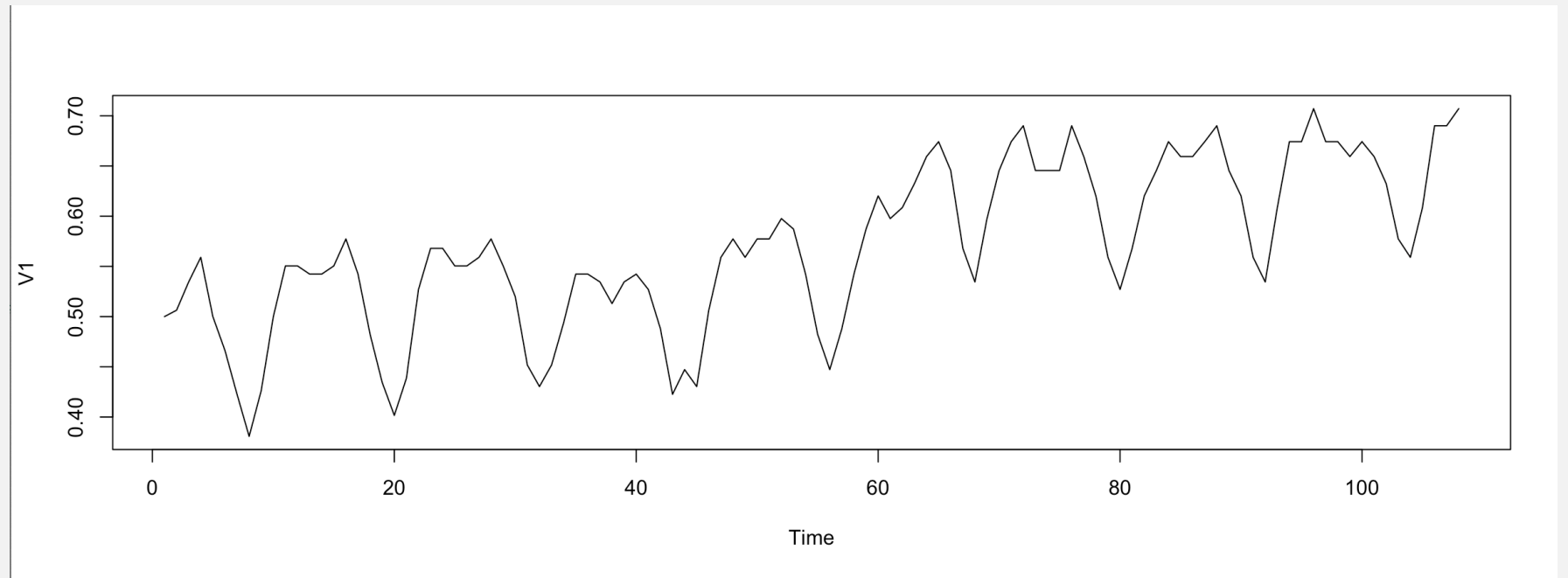
```
par(mfrow=c(2,1))  
acf(finalts, main="ACF", lag=36)  
pacf(finalts, main="PACF", lag=36)
```



## 분산안정화

- 분산안정화를 위해 자료를  $1/\sqrt{\text{data}}$ 의 형태로 변환하였다. 그 결과 분산이 상당부분 안정화된 것을 볼 수 있다.

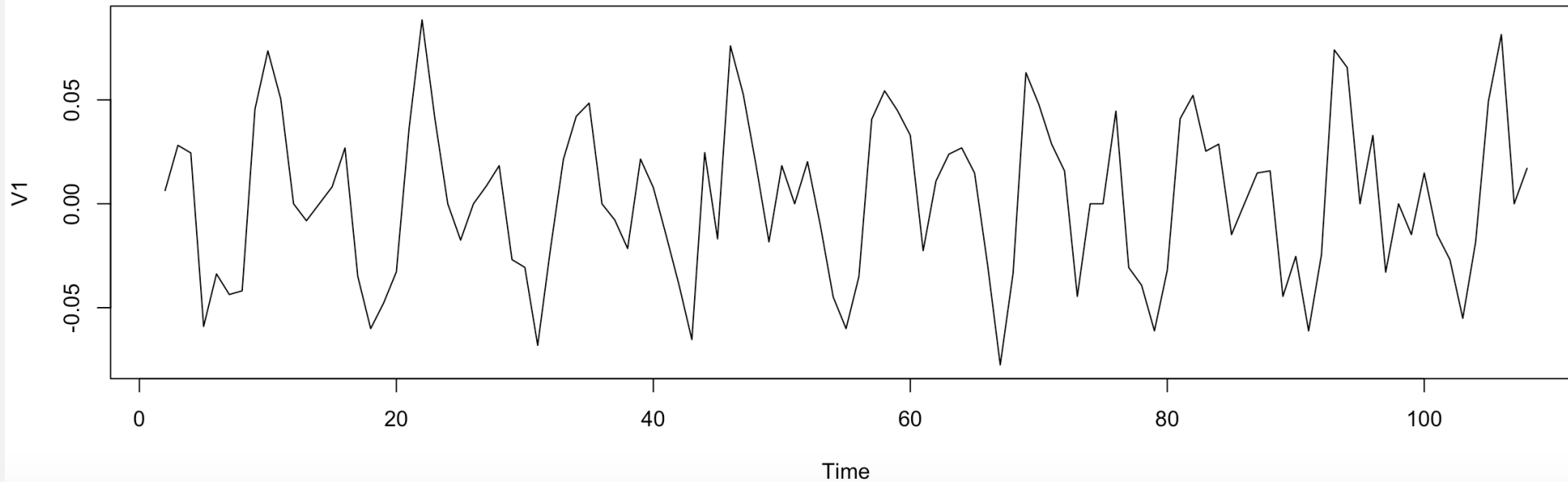
```
finalts_1<-1/sqrt(finalts)  
plot(finalts_1)
```



## 시계열 자료의 추세 제거(1차 차분)

- 1차 차분을 통해서 시계열 자료의 추세성을 없애기로 하였다. 그러자 다음 같은 시도표가 나왔고, 추세가 제거된 것을 확인할 수 있었다.

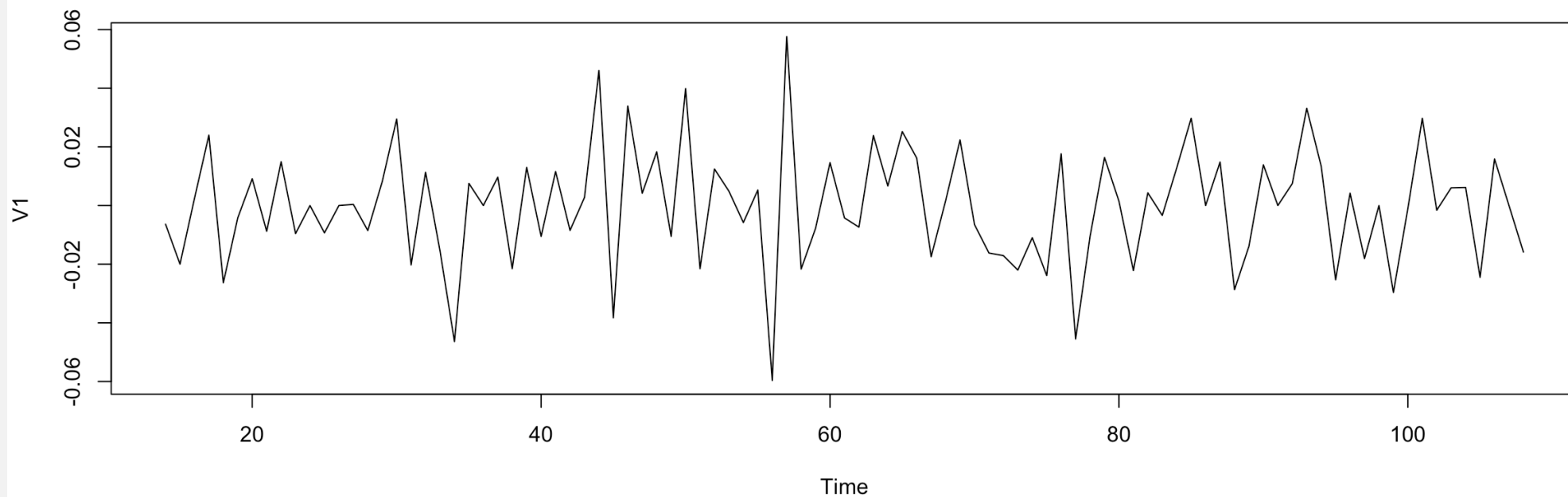
```
finalts_2<-diff(finalts_1)  
plot(finalts_2)
```



## 시계열 자료의 계절성 제거(2차 차분)

- 12 정도를 주기로 계절성이 존재하는 듯하여 2차 차분을 통해서 시계열 자료의 계절성을 없애기로 하였다. 그러자 다음 같은 시도표가 나왔고, 계절성이 제거된 것을 확인할 수 있었다.

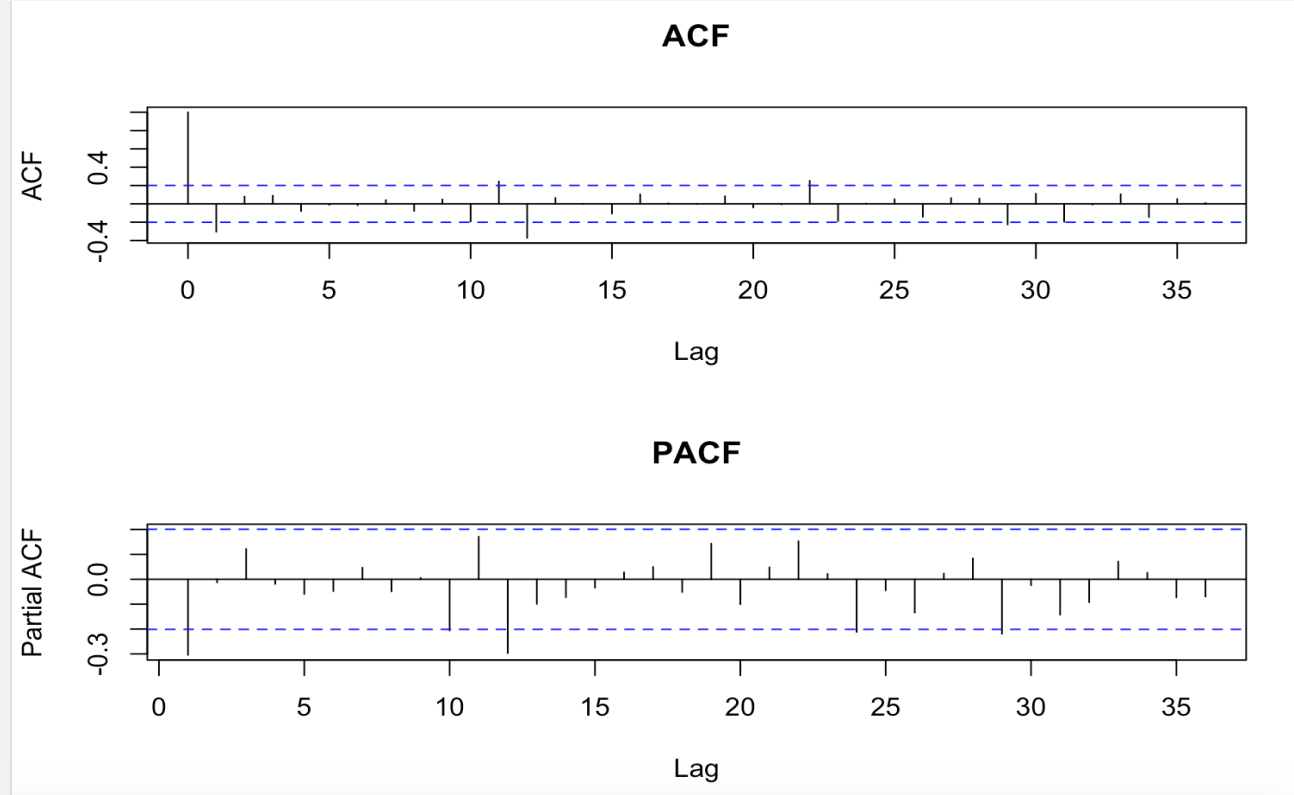
```
finalts_3<-diff(finalts_2, lag=12)  
plot(finalts_3)
```



## 조정된 시계열 자료의 ACF, PACF 확인

- 조정된 시계열의 ACF와 PACF를 확인해보니 다음과 같이 나왔다.  
변환자료를 살펴보면 다음과 같이 나온다. 해당 자료를 보면 확실하지는 않지만 둘다 지수적으로 감소한다고 가정하면 ARMA(2,1)이 무난하지 않을까 생각된다.

```
par(mfrow=c(2,1))  
acf(finalts_3, main="ACF", lag=36)  
pacf(finalts_3, main="PACF", lag=36)
```



## 시계열 모형 예상

실제로 `auto.arima` 코드를 이용하여 모형 차수를 예상해보니  $ARMA(2,1)$ 이 제시되었다. 그리하여 이를 기초로 하여 시계열 모형을 결정하였다.

```
Series: finalts_3
ARIMA(2,0,1) with zero mean

Coefficients:
            ar1      ar2      ma1
        -1.2425  -0.3574   0.9626
s.e.      0.0996   0.0960   0.0398

sigma^2 estimated as 0.0003513:  log likelihood=244.18
AIC=-480.36   AICc=-479.92   BIC=-470.15
```



## 시계열 자료의 모수 추정

그리하여 arima 코드를 이용하여 모수를 추정해보니 다음과 같은 결과가 나오게 되었다.

따라서 최종 모델은 다음과 같이 결정되었다.

$$1/\sqrt{X_t} \sim \text{SARIMA}(2,1,1) \times (0,1,0)_{12}$$

( $1/\sqrt{X_t}$ )로 표현된 이유는 해당 원 시계열이 현재 분산안정화를 위해  $1/\sqrt{\text{data}}$ 로 변환되었기 때문이다.)

```
fit<-arima(finalts_1, c(2,1,1), seasonal=list(order=c(0,1,0), period=12))
fit
```

Call:

```
arima(x = finalts_1, order = c(2, 1, 1), seasonal = list(order = c(0, 1, 0),
  period = 12))
```

Coefficients:

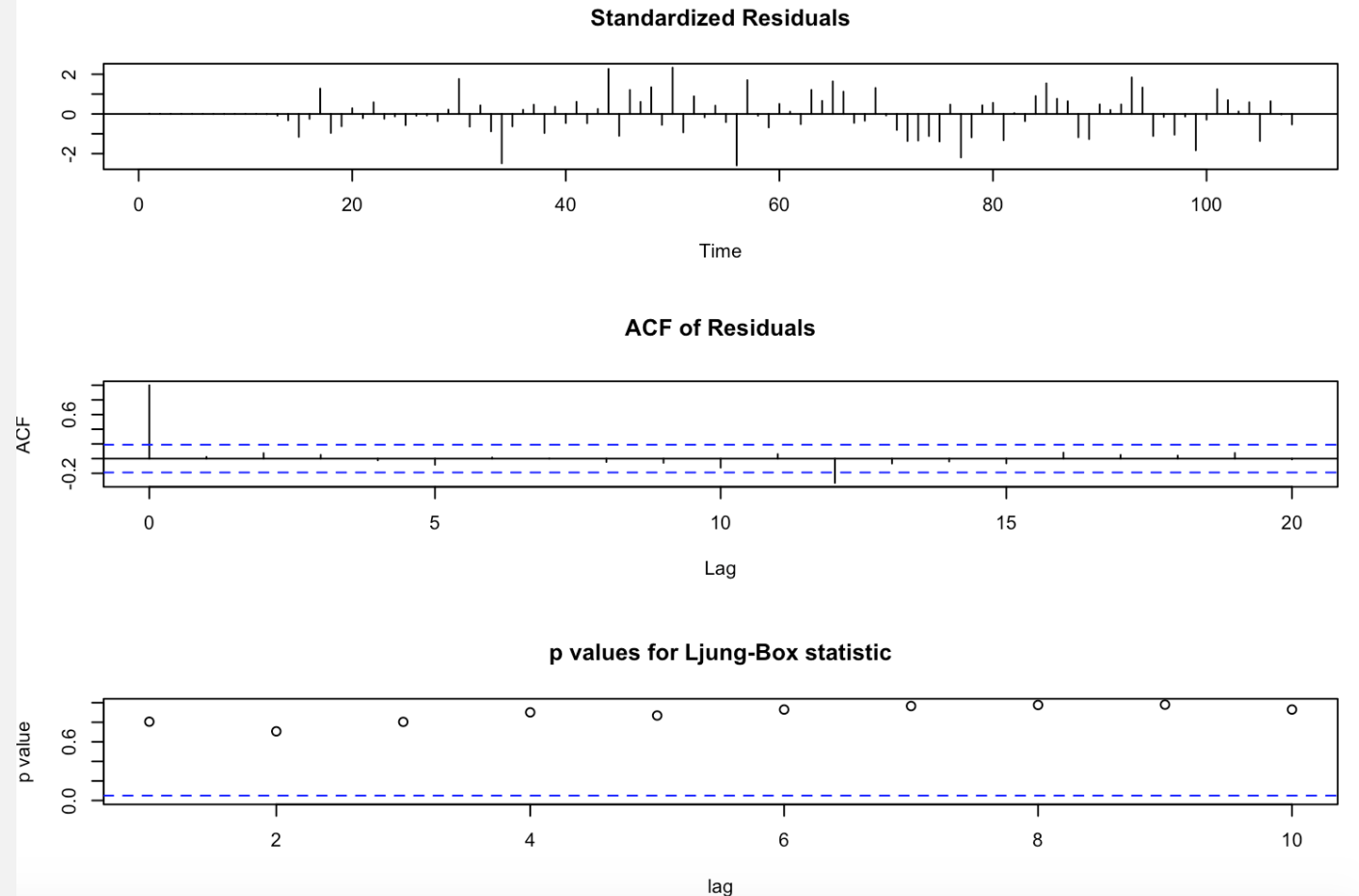
	ar1	ar2	ma1
	-1.2422	-0.357	0.9628
s.e.	0.0996	0.096	0.0397

sigma^2 estimated as 0.0003402: log likelihood = 244.18, aic = -480.36

## 시계열 모형에 대한 진단

- 모형이 자기 상관성, 정규성과 등분산성 가정을 만족하고 있는지를 확인하기 위해 `tsdiag` 코드를 이용하여 확인하였다. 그러자 다음과 같이 나왔고, 해당 가정들을 잘 만족하고 있는 것을 확인하였다.

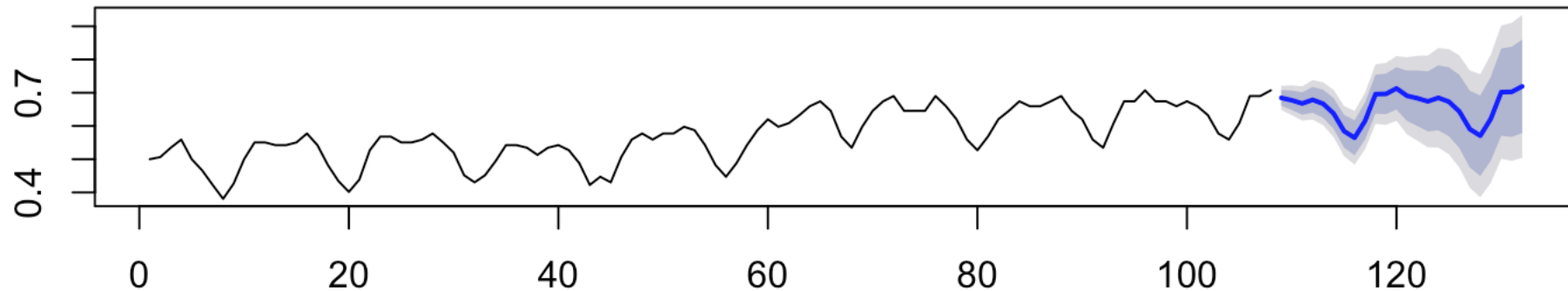
```
tsdiag(fit)
```



## 확정된 최종 시계열 모형 예측

- 이제 확정된 최종 모형을 구한 후, 변환된 데이터에 대해서 forecast 코드를 이용하여 예측값 12개(아직 원 데이터가 아님)를 생성해보았고, 이는 다음과 같았다. (사용한 코드: `forecast(fit)` , `plot(forecast(fit))` )

**Forecasts from ARIMA(2,1,1)(0,1,0)[12]**



## 시계열 데이터의 예측값

- predict 코드를 이용하여 예측값을 구하면 다음과 같은 예측값들이 나오게 된다 (아직 원 데이터의 예측값은 아님.)

```
> pred<-predict(fit, n.ahead=12)
> pred
$pred
Time Series:
Start = 109
End = 120
Frequency = 1
 [1] 0.6844920 0.6773704 0.6677237 0.6786596 0.6668179 0.6375803 0.5842849 0.5645289 0.6152138 0.6958138 0.6965129
[12] 0.7130016

$se
Time Series:
Start = 109
End = 120
Frequency = 1
 [1] 0.01844509 0.02273507 0.02624380 0.03031538 0.03285805 0.03613806 0.03840278 0.04111811 0.04321953 0.04556218
[11] 0.04753075 0.04961804
```

## 원 시계열 데이터의 예측값

- 그리고 원 데이터의 시계열 예측값을 구하게 되면, 다음과 같은 12개의 예측값이 나오게 된다.

```
> pred$pred^-2
```

```
Time Series:
```

```
Start = 109
```

```
End = 120
```

```
Frequency = 1
```

```
[1] 2.134338 2.179453 2.242882 2.171181 2.248980 2.459972 2.929211 3.137817 2.642091 2.065447 2.061302 1.967066
```

## 시계열 자료에 대한 예측값

- 이제 예측값을 토대로 plot를 그려보면 다음과 같이 나오게 된다.

```
ts.plot(finalts, pred$pred^-2, lty=c(1,3))
```

