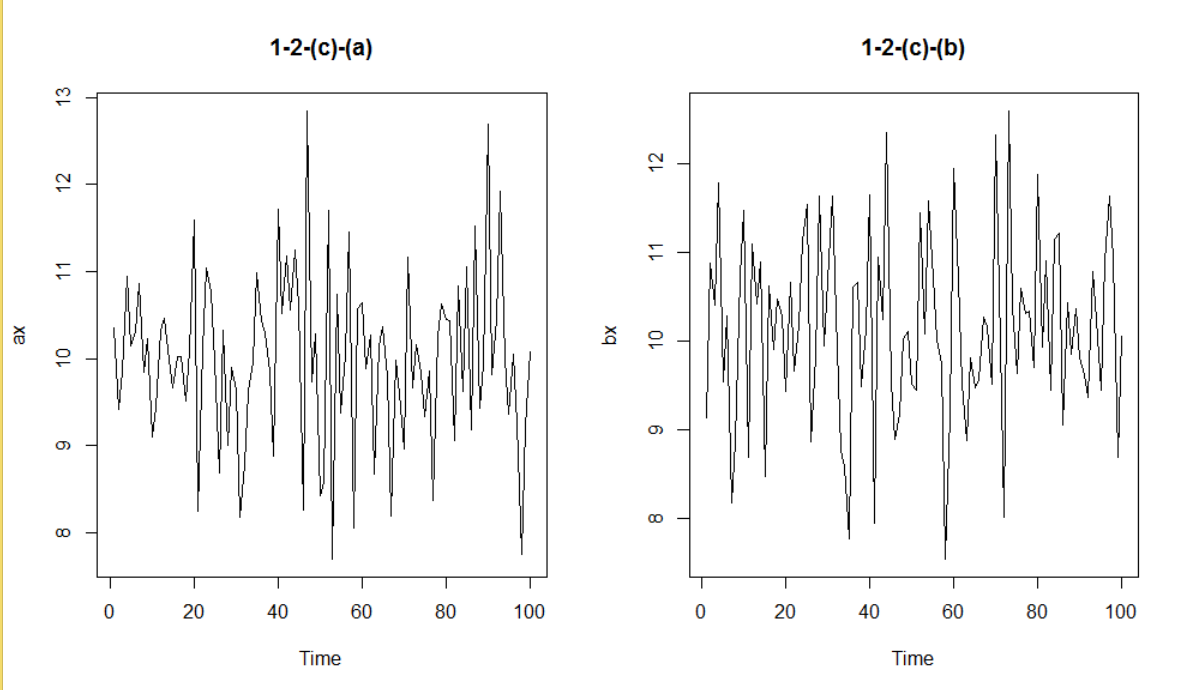
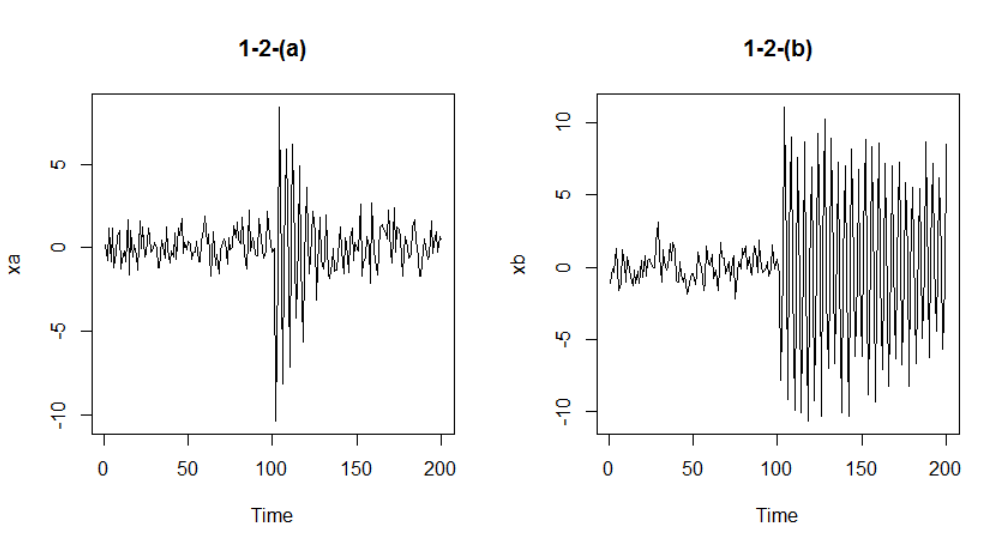
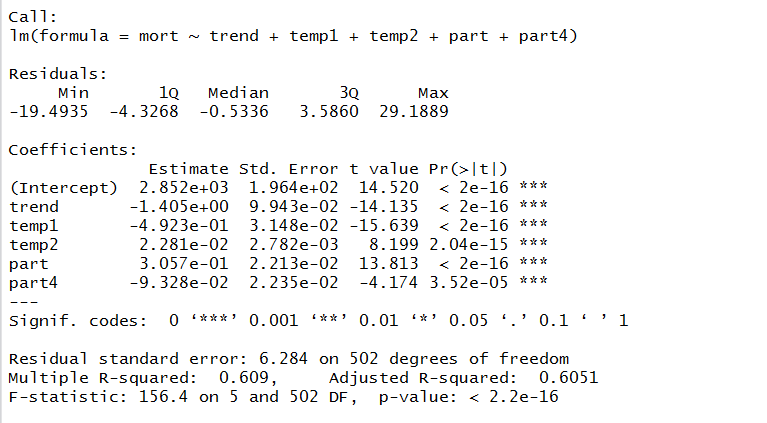
**A)**

**A-1) 1.2**



**A-2) 2.2**

(a)

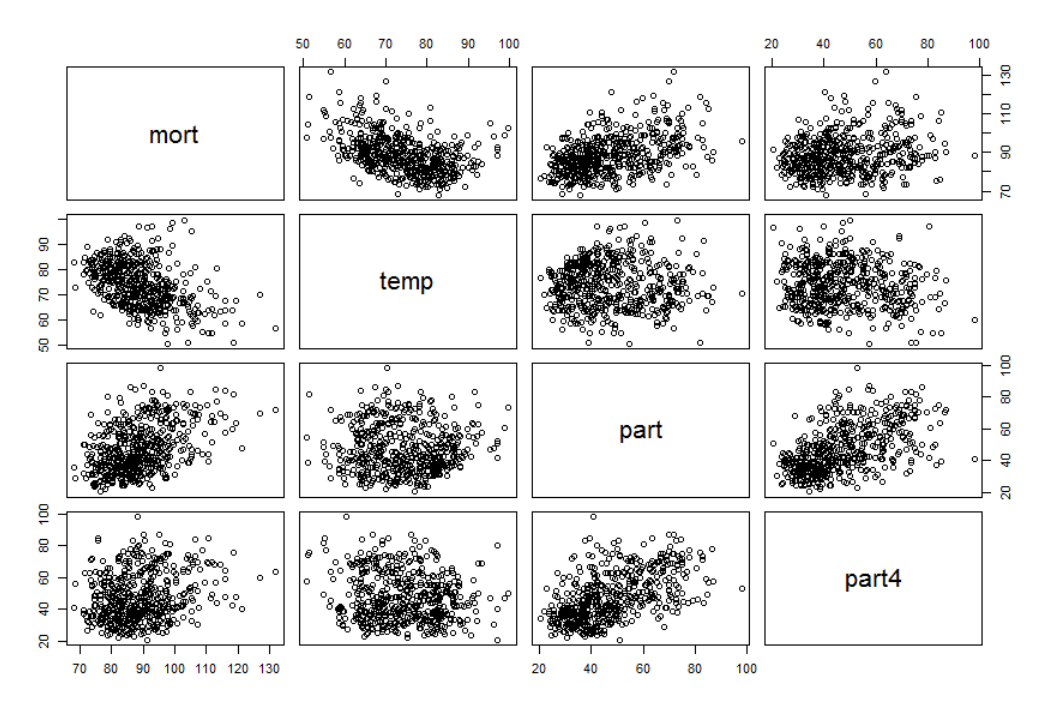


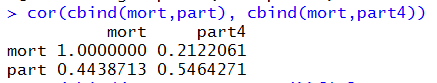
ANOVA table에서 p –value 검정결과 회귀식이 유의함.

각 변수들 p-value 검정결과 변수들이 잘 적합되었음을 확인할 수 있다.

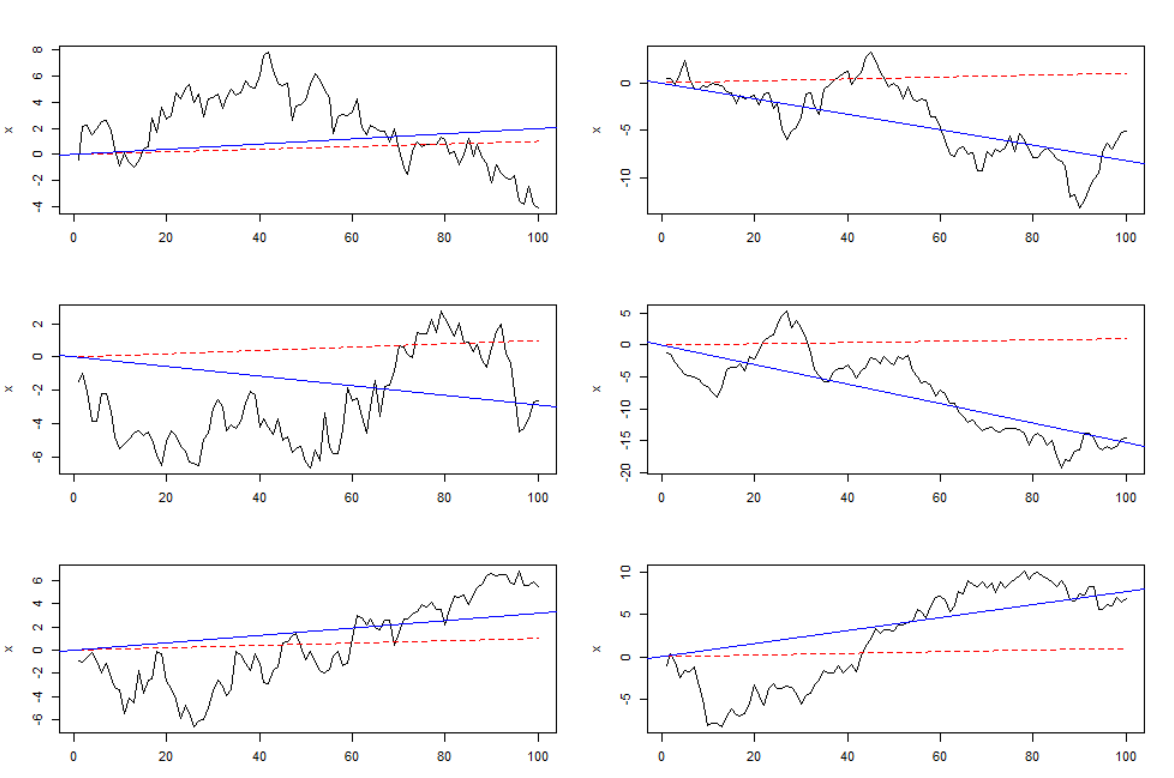
Adjusted R^2를 봤을 때 60%정도의 설명력을 지니는 것을 확인할 수 있다.

**(b)**



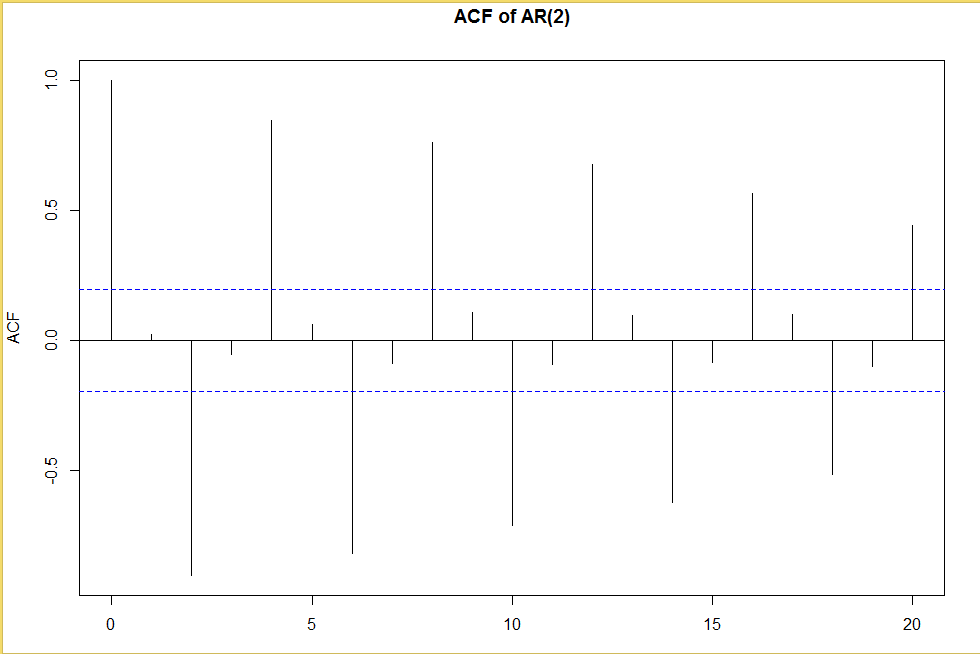


**A-3) 2-3**



Discuss my result : 평균(빨간선) 회귀선(파랑) / 평균은 점차 증가하는 추세이고, 이 추세는 함수 반복 횟수에 상관없이 동일 한데 (rnorm 에서 평균을 0.01로 지정해 줬으니 어쩌면 당연한 결과) 회귀선은 샘플이 어떻게 나타나느냐에 따라 굉장히 민감하게 반응함(매번 달라진다)

**A-4) 3.6**

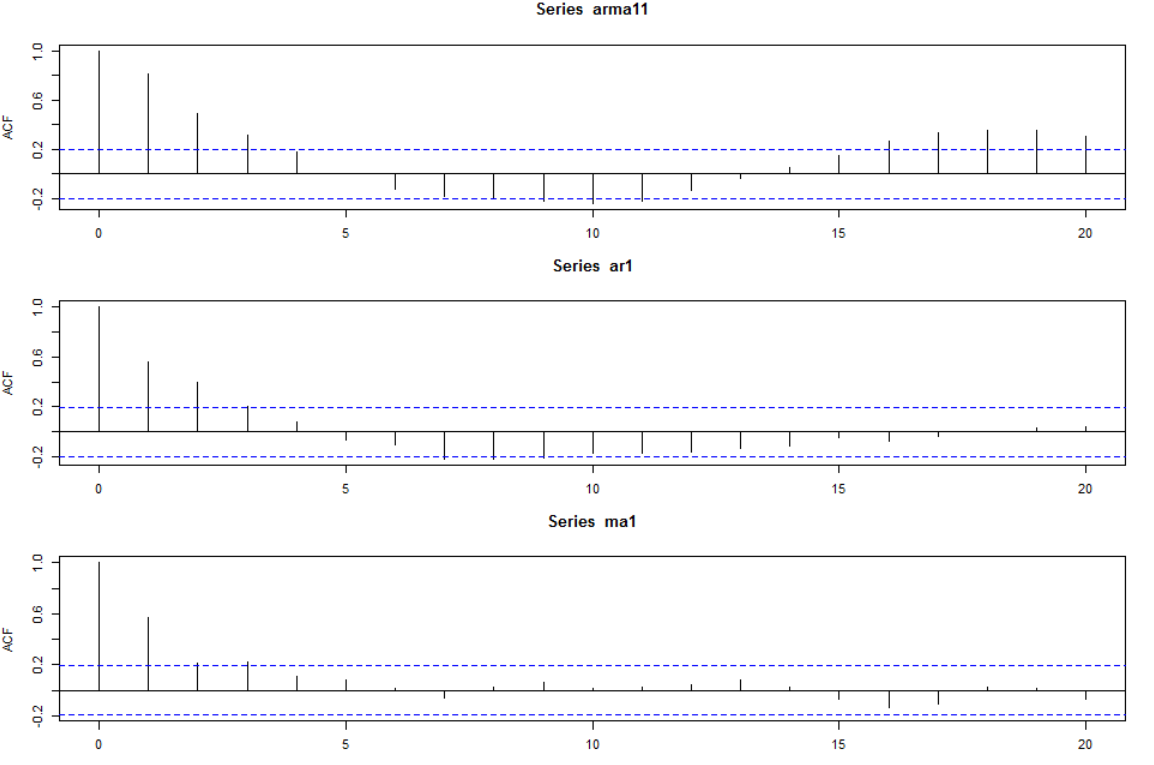


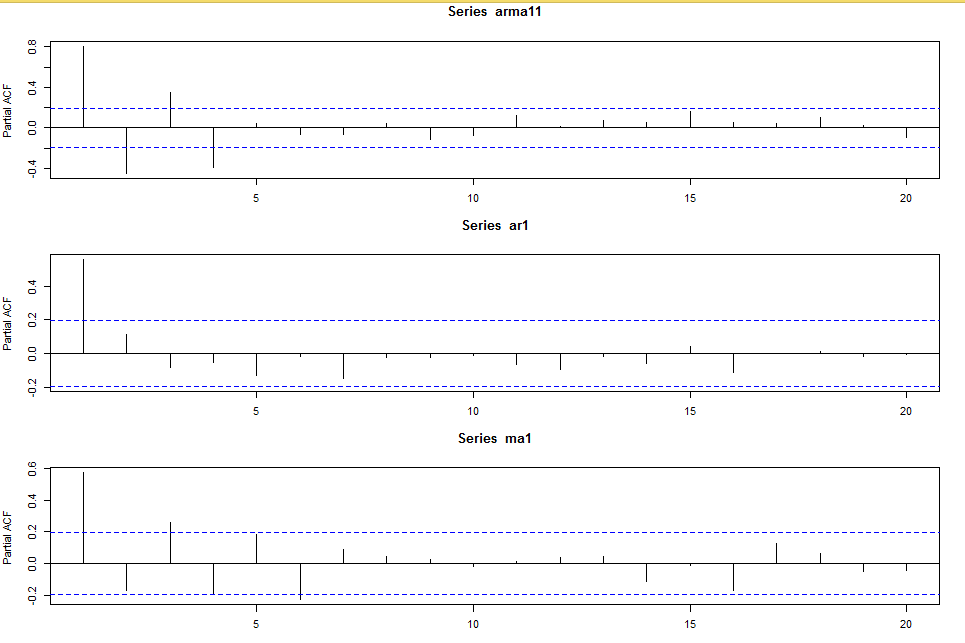
> polyroot(aa)[1]

[1] 1.054093+0i

현재시점의 데이터는 짝수번째 시차에 대한 데이터의 영향을 받고있다.

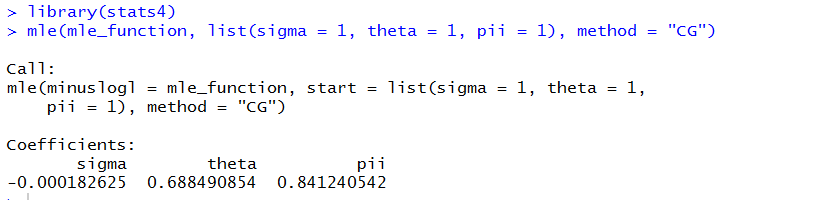
**A-5) 3-9**



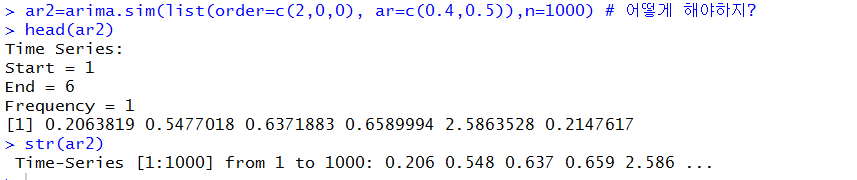


A-6)3-19

**A-7)3-21**

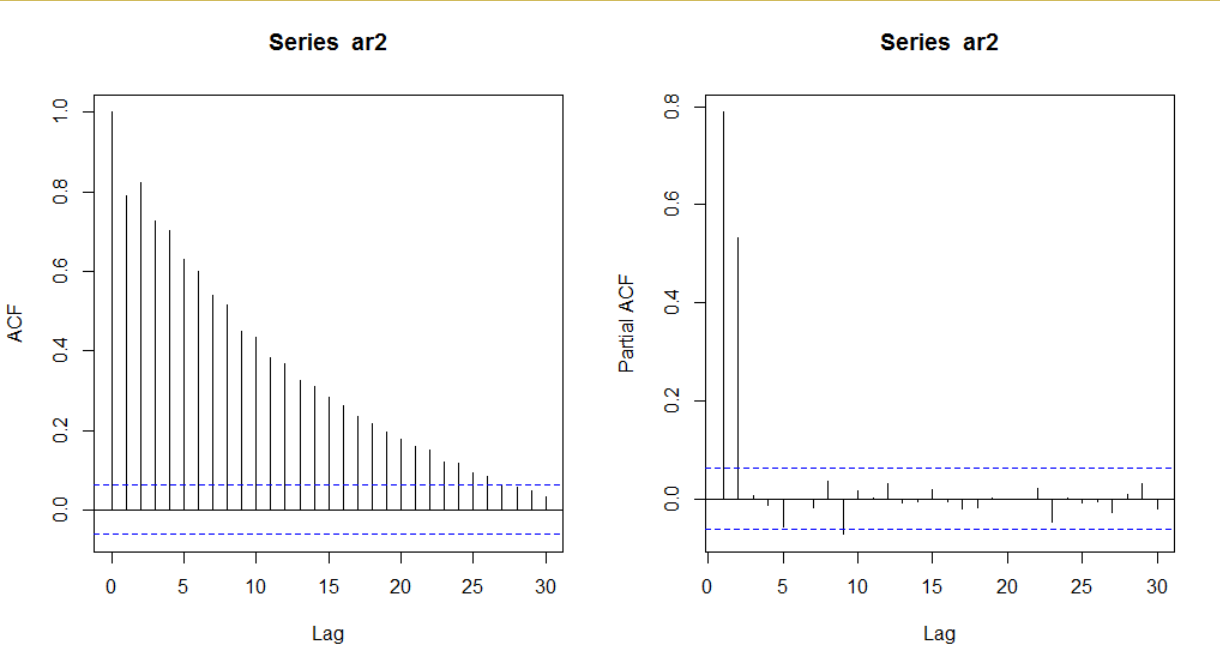


**B) Generating 1000 observations form AR(2) model**

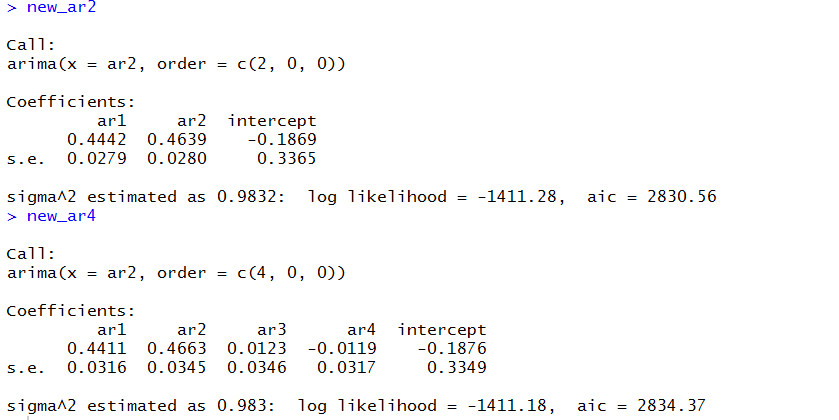


Xt=0.4Xt-1-0.5Xt-2+Zt ( Zt~WN) 인 임의의 모델 생성

**B-i) Draw time series plot, sample ACF and PACF plot.**

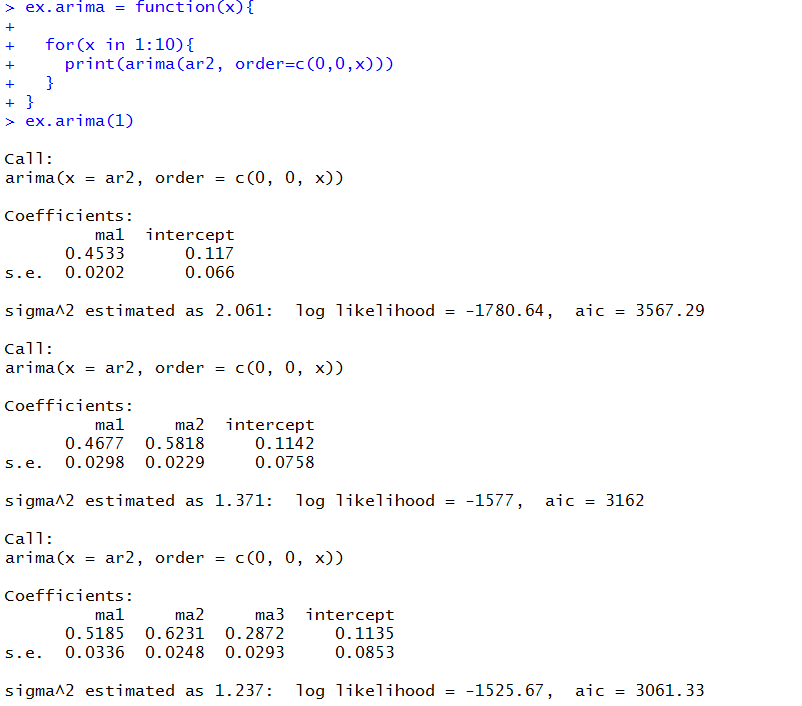


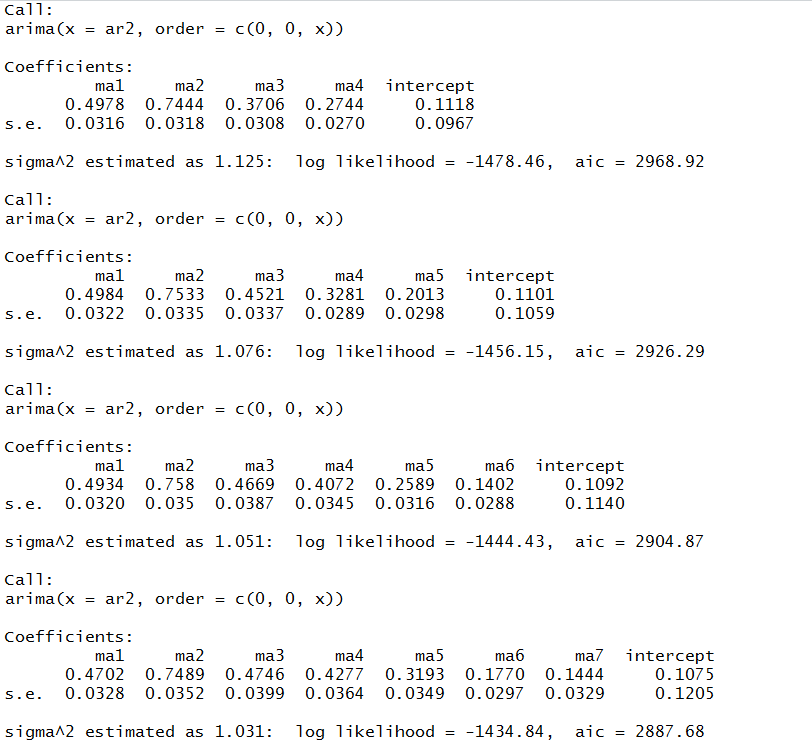
**B-ii) Fit the model under AR(2) and under AR(4) and explain the result. What do you think is the problem when you fit bigger model then necessary?**

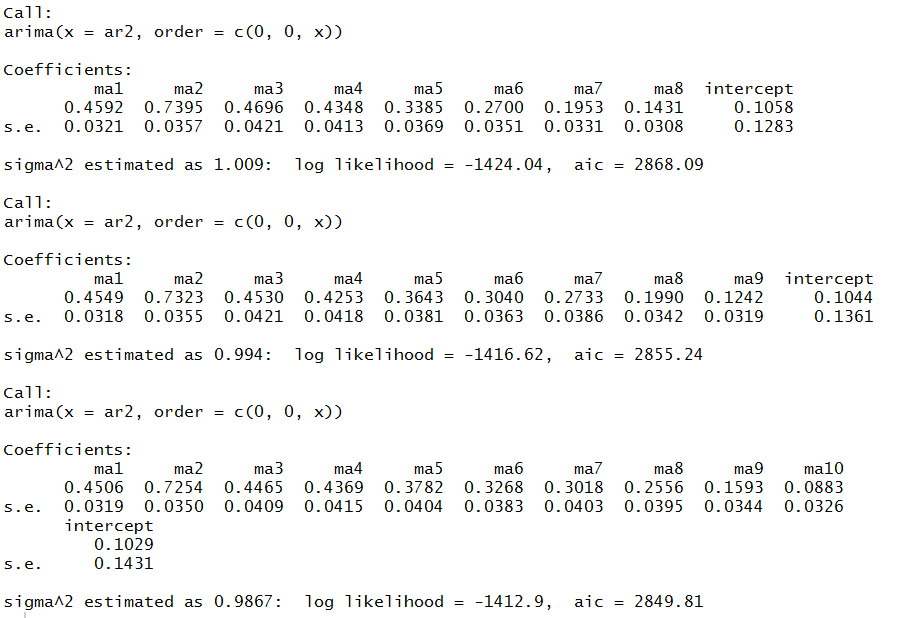


문제점: coefficients 값을 보아 ar3, ar4이 유의하지도 않을 뿐더러 AIC값 역시 커졌다. 이를 통해AR(2)가 더 적합하다는 걸 파악할 수 있다. (유의성은 ar의 계수와 s.e.의 값을 통해 알 수 있다)

**B-iii) Fit the model under MA(p) varying p from 1 to 10. What do you observe? Explain your observation**





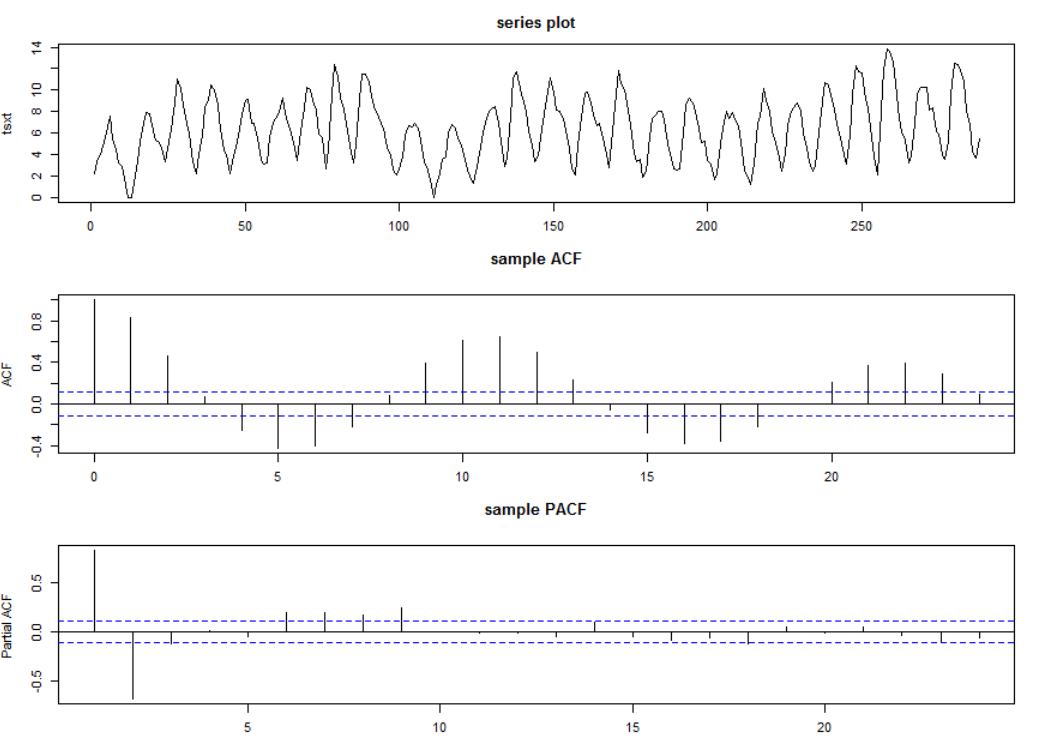


모형이 커질수록 정확도는 높아진다. (log likelihood 값과 AIC값 모두가 감소함을 알 수 있다)

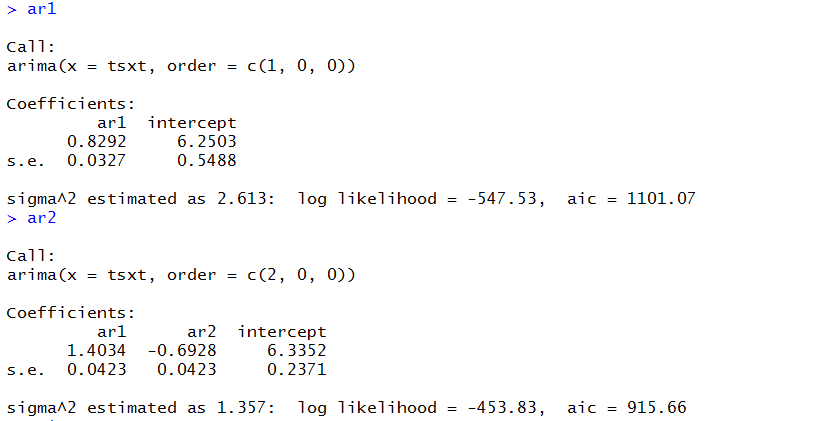
분산 역시 작아진다. 하지만 AIC가 낮아진다고 무조건적으로 모형이 큰 게 좋은게 아니므로 적당한 기준을 세워야 할 것이다. 보면 점점 MA값이 커질수록 설명력이 낮아지는데 이를통해 MA(4)정도까진 선택해도 될 것 같다.

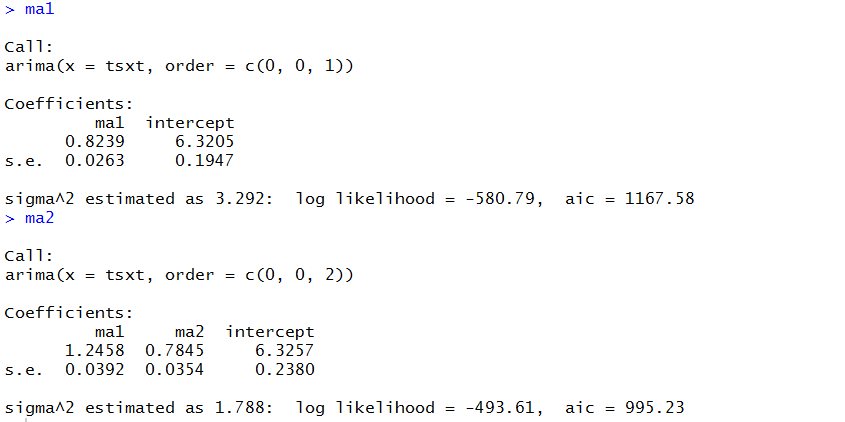
**C)**

**C-i) Draw time series plot, sample ACF and PACF plot**



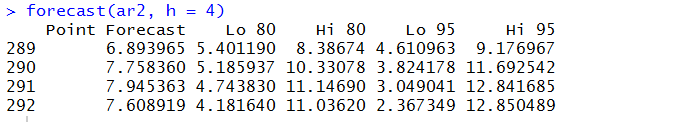
**C-ii) By considering the sample ACF and sample PACF, decide wich of the following would e appropriate for this data: AR(1), AR(2), MA(1), MA(2). Use the data to estimate the parameters of the model that you choose.**



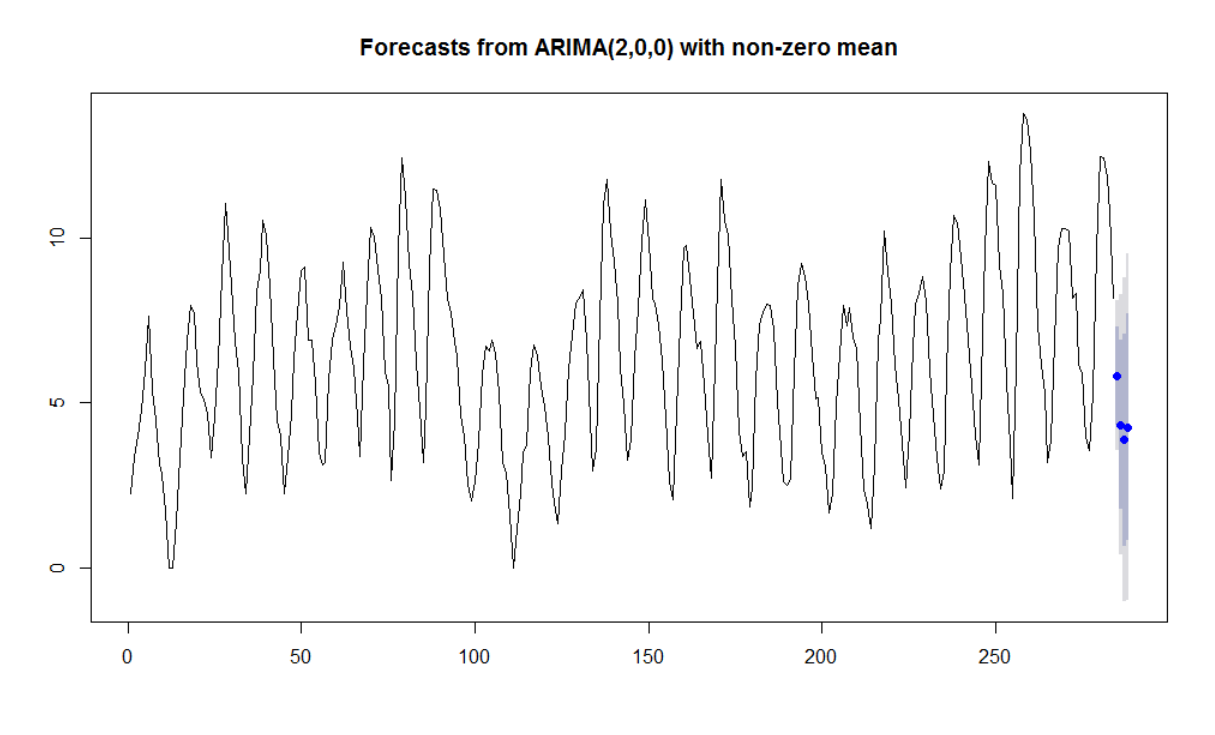


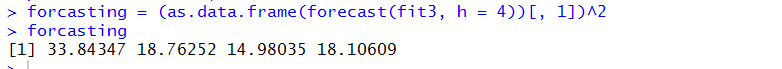
Conclusion : AIC가 제일 낮은 AR(2)가 제일 적합하다. Xt=1.4034\*Xt-1-0.6928\*Xt-2+Zt (Zt~WN)

**C-iii) Using your fitted model, calculate forecasts Xn^n+h for h=1,2,3,4. Calculate the 95% prediction intervals (assuming Gussian noise)**

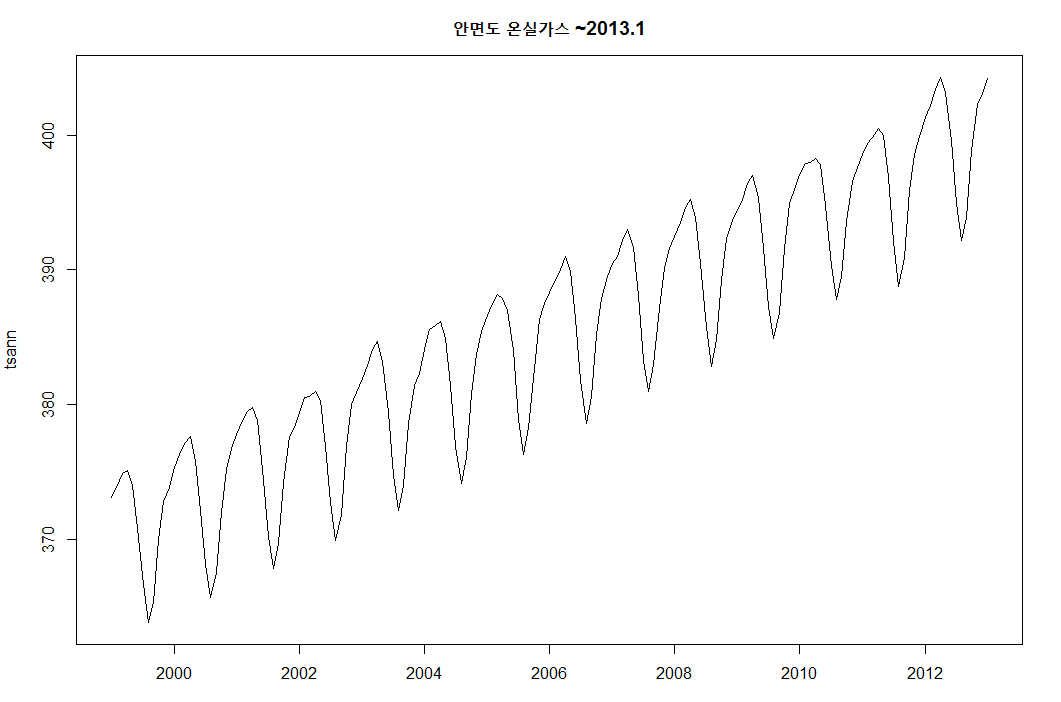


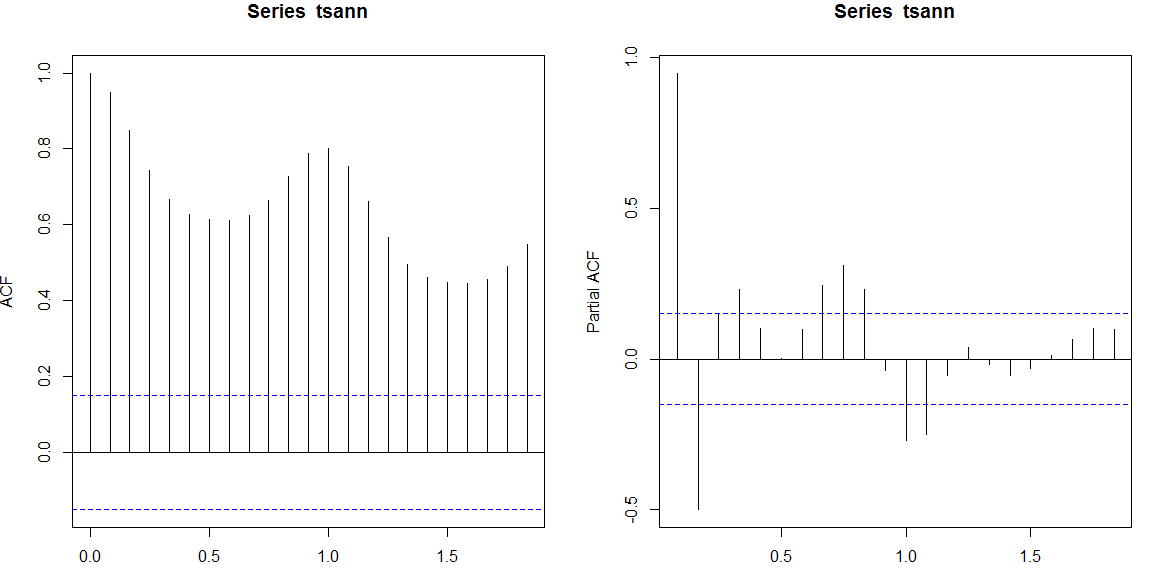
**C-iv) Use the observation from 1700 to 1983 for fitting the model. Plot all of the data, and your forecasts and prediction intervals for the last four years. (Don’t forget to undo the square root transformation by taking the square of your predictions)**



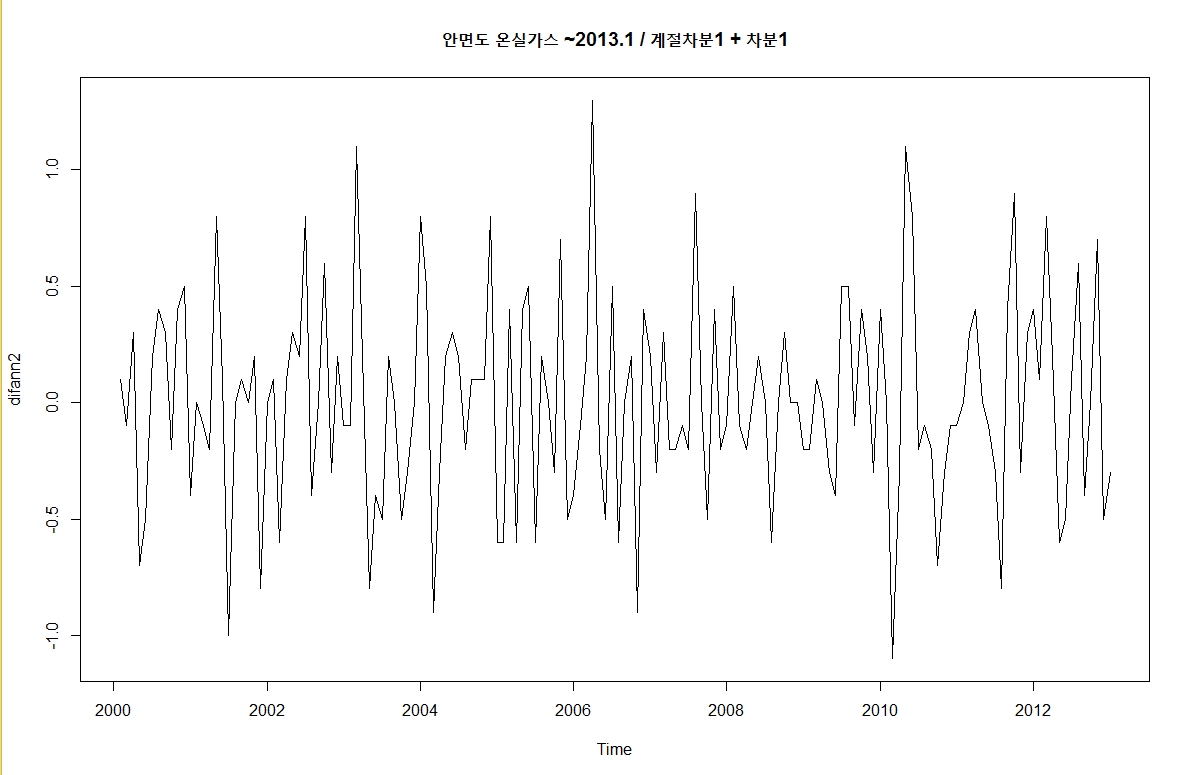


**D) By using seasonal ARIMA models. Forecast monthly averages for year 2014. Compare your forecast with the observed values. Do anything you can and make a readable answer.**

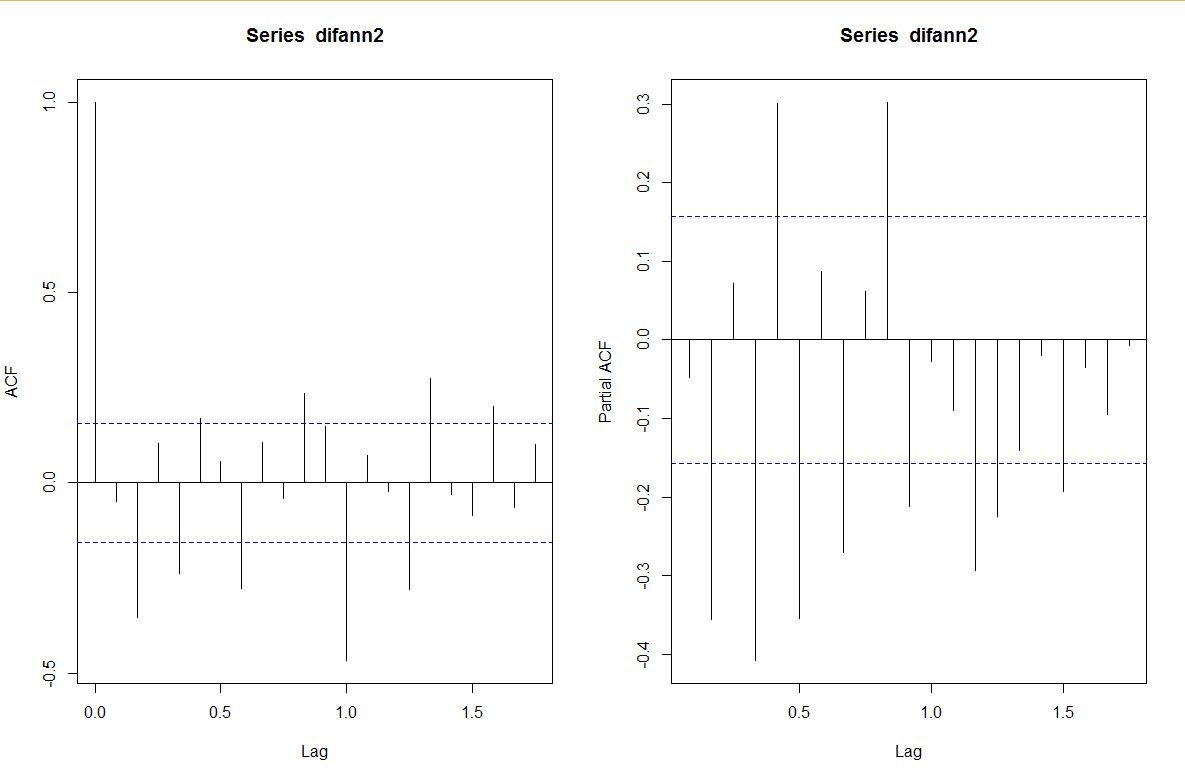




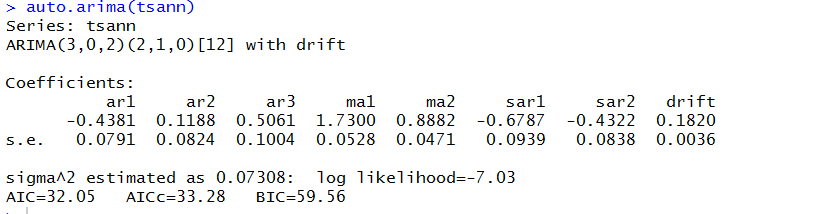
Trend , 주기성 존재 -> differncing 으로 주기와 트렌드 제거하는 것이 필요하다



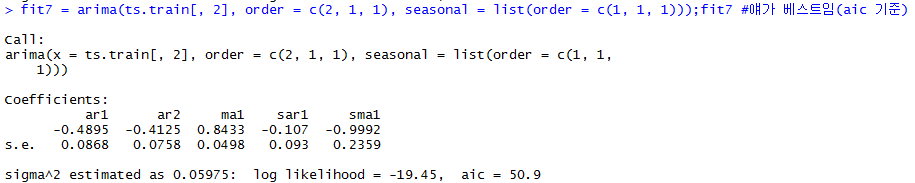
트렌드 제거, 주기 제거 확인. (계절차분 한번 + 차분한번)



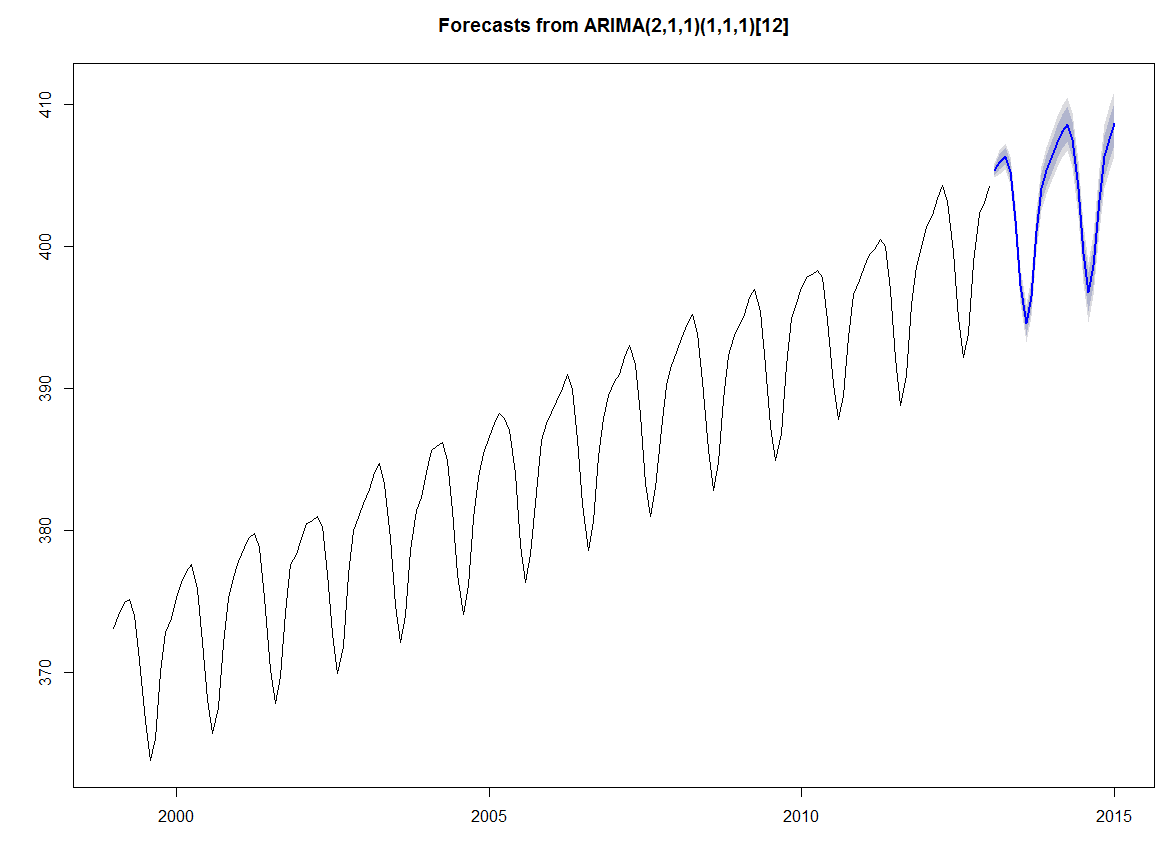
ARMA(4,1) 정도가 적합하다고 일단은 판단함.



이를 기반으로 여러가지를 fitting 해본 결과



Arima(2,1,1)\*(1,1,1)[12] 가 AIC기준으로 가장 좋은 적합모델임을 확인했다.





결과값이 별차이가 없다!!!!!