Stochastic Modeling Assignment 4

Mengting Ding, Vanessa Guzman, Zouzi Qi, Nash Kleisner

12/3/2021

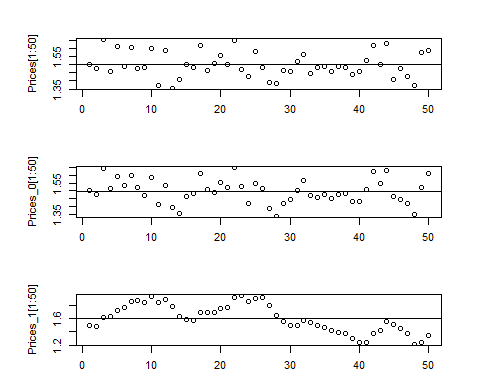
rm(list=ls())  
library(dplyr)  
library(ggplot2)

## 1)

PStar <- 1.5  
Phi <- -0.4  
Prices <- c(1.5)  
set.seed(13)  
et <- c()  
for (i in 1:10000){  
 shock <- rnorm(1,0,0.08)  
 et <- c(et,shock)  
 Prices <- c(Prices,PStar\*(1-Phi)+Phi\*Prices[i-1]+shock)}  
  
Phi\_0 <- 0  
Prices\_0 <- c(1.5)  
set.seed(13)  
for (i in 1:10000){  
 shock <- rnorm(1,0,0.08)  
 Prices\_0 <- c(Prices\_0,PStar\*(1-Phi\_0)+Phi\_0\*Prices\_0[i-1]+shock)}  
  
Phi\_1 <- 1  
Prices\_1 <- c(1.5)  
set.seed(13)  
et <- c()  
for (i in 1:10000){  
 shock <- rnorm(1,0,0.08)  
 et <- c(et,shock)  
 Prices\_1 <- c(Prices\_1,PStar\*(1-Phi\_1)+Phi\_1\*Prices\_1[i-1]+shock)}

## 2)

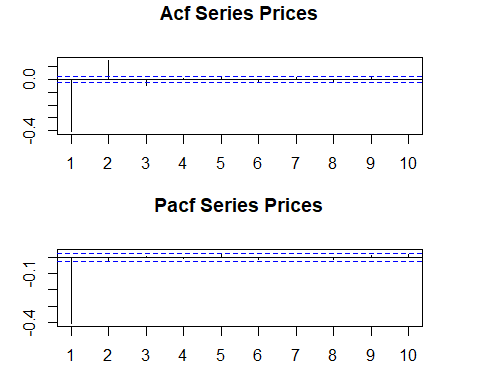
par(mar=c(3,6,3,3))  
par(mfrow=c(3,1))  
plot(Prices[1:50]) %>% abline(h=mean(Prices[1:50]))  
plot(Prices\_0[1:50]) %>% abline(h=mean(Prices\_0[1:50]))  
plot(Prices\_1[1:50]) %>% abline(h=mean(Prices\_1[1:50]))



Our data series looks stationary as the data points are within fairly equal distances from the mean. Phi equals to 0 has no noise filtered as P star is in full effect. Phi equals to 1 has more noise filtered out as there is no multiplier effect with P star, thus it is more effective in predicting prices.

## 3)

library(forecast)  
par(mar=c(3,3,3,3))  
par(mfrow=c(2,1))  
Acf(Prices,lag.max=10,plot=TRUE,type="correlation", ci=0.99, main="Acf Series Prices")  
Pacf(Prices,lag.max=10,plot=TRUE, ci=0.99, main="Pacf Series Prices")



#### Statistically significant

Acf has 3 statistically significant correlations. Pacf has 1 statistically significant correlation.

#### Decaying behavior

Acf shows an exponential decay to zero as the lag increases. Pacf shows a rapid decay to zero as the lag increases.

#### Correlations alternating in sign

Both Acf and Pacf oscillate indicating that the correlations alternate in sign.

## 4)

library(dynlm)  
series <- ts(Prices, frequency=12)  
options(scipen=9999)  
dynlm(Prices~L(Prices,-1))

##   
## Time series regression with "numeric" data:  
## Start = 1, End = 10000  
##   
## Call:  
## dynlm(formula = Prices ~ L(Prices, -1))  
##   
## Coefficients:  
## (Intercept) L(Prices, -1)   
## -0.000000000000003126 1.000000000000001998

Intercept: -0.000000000000003126  
Slope: 1

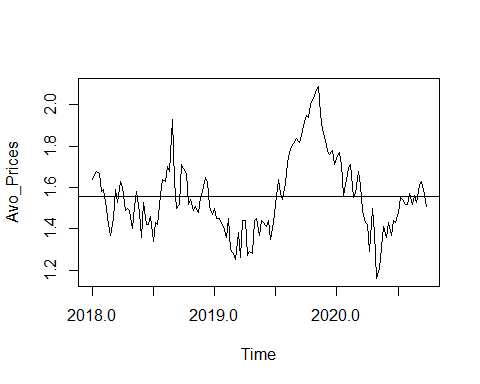
## 5)

avocado <- read.csv("avocado(5).csv")  
set.seed(10)  
Avo\_Prices <- ts(avocado$Avocado\_Prices,start = c(2018,1),frequency = 52)  
head(Avo\_Prices)

## Time Series:  
## Start = c(2018, 1)   
## End = c(2018, 6)   
## Frequency = 52   
## [1] 1.64 1.66 1.68 1.67 1.58 1.59

## 6)

p <- plot(Avo\_Prices) %>% abline(h=mean(Avo\_Prices))



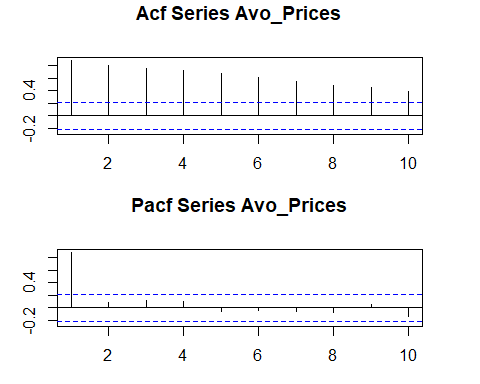
mean(Avo\_Prices)

## [1] 1.557413

The time mean of the series is 1.557413.

## 7)

library(forecast)  
par(mar=c(3,3,3,3))  
par(mfrow=c(2,1))  
Acf(Avo\_Prices, plot=TRUE, lag.max=10, ci=0.99, main="Acf Series Avo\_Prices")  
Pacf(Avo\_Prices, plot=TRUE, lag.max=10, ci=0.99, main="Pacf Series Avo\_Prices")



#### Statistically significant

All of the correlations of Acf here are statistically significant here whereas there were only 3 statistically significant correlations in 3). Both Pacfs in 3) and here only have 1 statistically significant correlation.

#### Decaying behavior

Acf 3) decays faster than Acf decays here. Pacf in 3) seems to decay just as quickly as Pacf here.

#### Correlations alternating in sign

Acf correlations in 3) alternated in sign but Acf correlations here are all positive (therefore do not alternate in sign). The single statistically significant correlation in Pacf in 3) is negative whereas the single statistically significant correlation in Pacf in 7) is positive.

## 8)

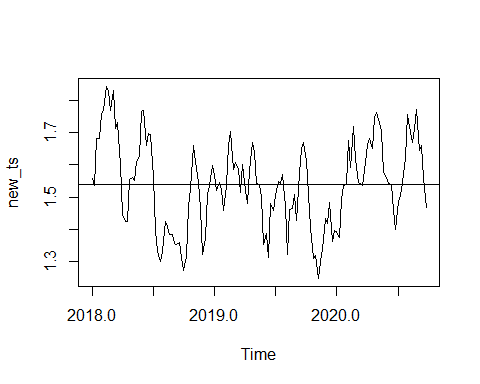
library(forecast)  
AR\_1 <- Arima(Avo\_Prices, order=c(1,0,0))  
AR\_1

## Series: Avo\_Prices   
## ARIMA(1,0,0) with non-zero mean   
##   
## Coefficients:  
## ar1 mean  
## 0.8822 1.5591  
## s.e. 0.0376 0.0564  
##   
## sigma^2 estimated as 0.007081: log likelihood=151.29  
## AIC=-296.59 AICc=-296.42 BIC=-287.7

The suggested value of Phi is 0.8822. The implied long run equilibrium price is $1.5591.

## 9)

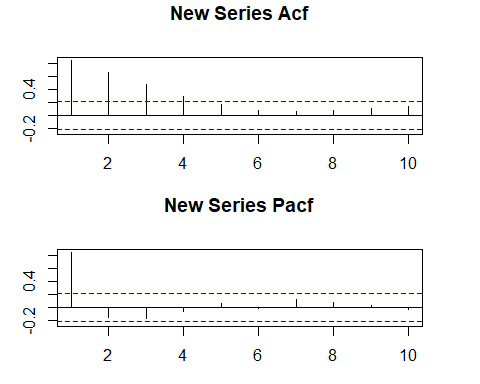
PStar <- 1.5591  
Phi <- 0.8822  
Simulated\_Prices <- c(1.5591)  
set.seed(13)  
et <- c()  
for (i in 1:143){  
 shock <- rnorm(1,0,0.08)  
 et <- c(et,shock)  
 Simulated\_Prices <- c(Simulated\_Prices,PStar\*(1-Phi)+Phi\*Simulated\_Prices[i-1]+shock)  
}  
new\_ts <- ts(Simulated\_Prices,frequency = 52,start=c(2018,1))  
plot(new\_ts) %>% abline(h=mean(new\_ts))



(mean(new\_ts))

## [1] 1.540201

par(mar=c(3,3,3,3))  
par(mfrow=c(2,1))  
Acf(new\_ts, lag.max=10, ci=0.99, main="New Series Acf")  
Pacf(new\_ts, lag.max=10, ci=0.99, main="New Series Pacf")

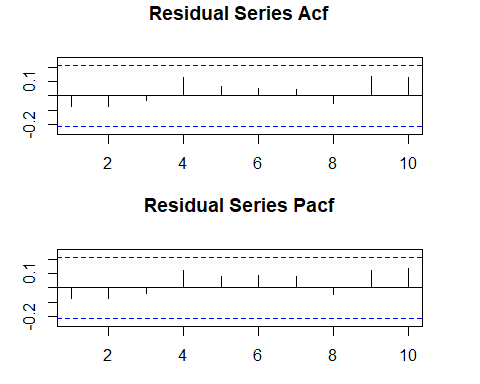


There are only 4 statistically significant correlations in Acf here whereas all of the correlations of Acf in 7) are statistically significant. However, all correlations of both Acfs are positive.

The ACF plot looks fairly the same here as it does in 7) where the first correlation of each is positive and significant and the rest of the correlations are oscillating and not significant.

## 10)

res <- AR\_1$residuals  
par(mar=c(3,3,3,3))  
par(mfrow=c(2,1))  
Acf(res,lag.max=10, ci=0.99, main="Residual Series Acf")  
Pacf(res,lag.max=10, ci=0.99, main="Residual Series Pacf")



#white noise process  
Box.test(res, type=c("Ljung"),lag=10)

##   
## Box-Ljung test  
##   
## data: res  
## X-squared = 12.003, df = 10, p-value = 0.2848

Though both plots have correlations alternating in sign, there are no significant lags in either of the plots.

ACF correlations in 7) are all statistically significant and ACF correlations here are all not statistically significant. ACF correlations in 7) are all positive and ACF correlations here alternate in sign.

PACF in 7) has 1 statistically significant correlation and PACF correlations here are all not statistically significant. PACF correlations both in 7) and here alternate in sign.

The p-value of 0.2848 is high, so we have to reject the null hypothesis, meaning that the model shows a lack of fit. There is no discernible pattern in the residuals, meaning that the residuals are random, which is what we would prefer to have.

## 11)

forecast1 <- forecast(AR\_1,h=5)  
#length(forecast1)  
forecast1

## Point Forecast Lo 80 Hi 80 Lo 95 Hi 95  
## 2020.750 1.515782 1.407940 1.623623 1.350852 1.680711  
## 2020.769 1.520882 1.377074 1.664690 1.300947 1.740817  
## 2020.788 1.525381 1.358875 1.691888 1.270732 1.780031  
## 2020.808 1.529351 1.347125 1.711576 1.250661 1.808041  
## 2020.827 1.532853 1.339275 1.726430 1.236801 1.828904

## 12)

plot(forecast1)

