Homework 5 Solution

March 9, 2019

Problem 5.2.

(a):$Y_t = 3 + Y_{t-1} + e_t - 0.75e_{t-1}$
Here we see that

$$\nabla Y_t = Y_t - Y_{t-1} = 3 + e_t - 0.75e_{t-1};$$
$$\text{Thus, } \mathbb{E}(\nabla Y_t) = 3 \text{ and } \mathrm{Var}(\nabla Y_t) = (1 + 0.75^2)\sigma_e^2 = \frac{25}{16}\sigma_e^2.$$

(b):$Y_t = 10 + 1.25Y_{t-1} - 0.25Y_{t-2} + e_t - 0.1e_{t-1}$
In this case:

$$\nabla Y_t = Y_t - Y_{t-1} = 10 + 0.25(Y_{t-1} - Y_{t-2}) + e_t + 0.1e_{t-1}.$$

thus we could see that the new model is stationary and invertible, which means that the original model is ARIMA(1,1,1) model with $\phi_1 = 0.25,\ \theta_1 = 0.1,\ \theta_0 = 10$. Then we could use the formulas that we learned before and get the following:

$$\mathbb{E}(\nabla Y_t) = \frac{\theta_0}{1 - \phi_1} = \frac{10}{1 - 0.25} = \frac{40}{3}.$$
$$\mathrm{Var}(\nabla Y_t) = \frac{1 - 2\phi_1\theta_1 + \theta_1^2}{1 - \phi_1^2}\sigma_e^2 = 1.024\sigma_e^2.$$

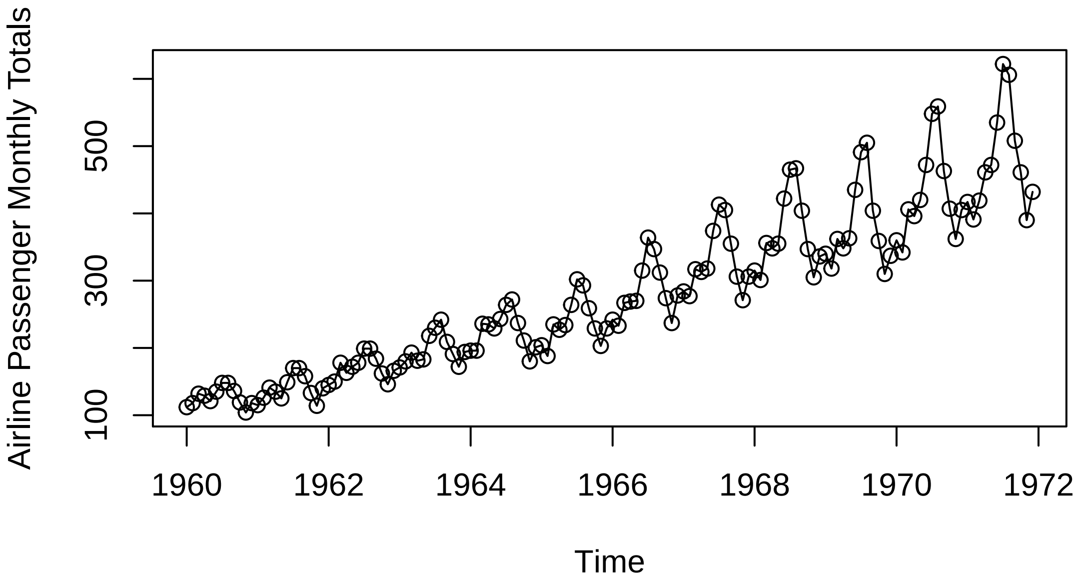
Problem 4.

(i):We know from the theorem that, for large n, r_k is approximately normally distributed with mean ρ_k , variance c_{kk}/n and $\mathrm{Corr}(r_k, r_j) \approx c_{kj}\sqrt{c_{kk}c_{jj}}$, where c_{ij} is given by (6.1.2). For a white noise process, it is obvious that the true autocorrelation $\rho_k = 0$ for all $k \neq 0$ and $\rho_0 = 1$. Then plug these values into (6.1.2) we could easily see that $c_{kk} = 1$. Combining all the information, we have that $r_k, k \geq 1$ is approximately distributed as N(0,1/n) for large n.

(ii): From the theorem, we know that $\sqrt{n}(r_1 - \rho_1), \sqrt{n}(r_2 - \rho_2), \cdots, \sqrt{n}(r_m - \rho_m)$ are approximately multivariate normal distributed when n is large, where $\mathrm{Corr}(r_k, r_j) \approx c_{kj}\sqrt{c_{kk}c_{jj}}$. It is easily seen that $c_{ij} = 0$ for all $i \neq j$, which means r_k and r_j are uncorrelated.

5.13 (a):

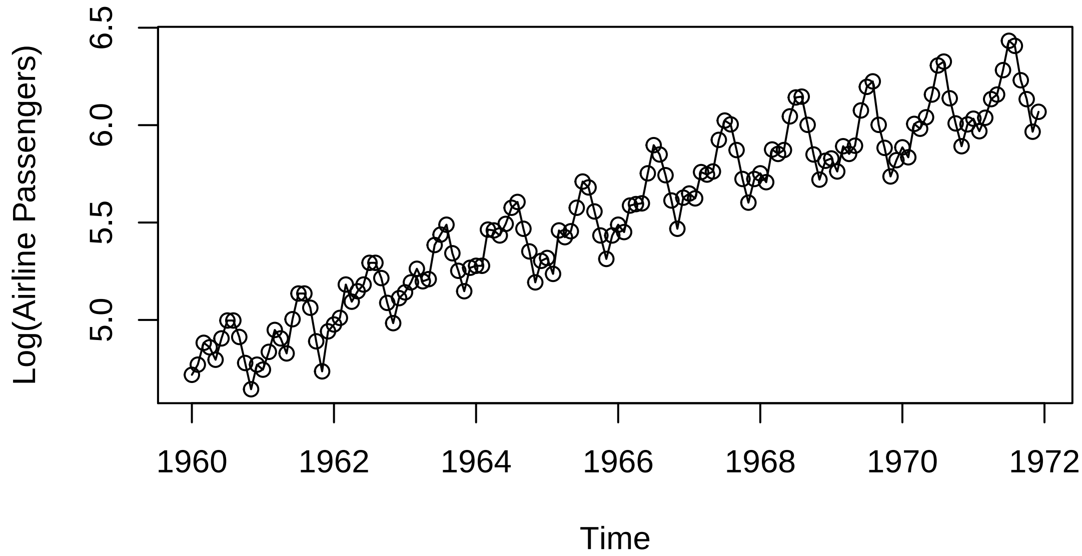
```r
data(airpass)
plot(airpass,type='o',ylab='Airline Passenger Monthly Totals')
```



There is a general upward "trend" with increased variation at the higher levels. There is also evidence of seasonality.

(b):

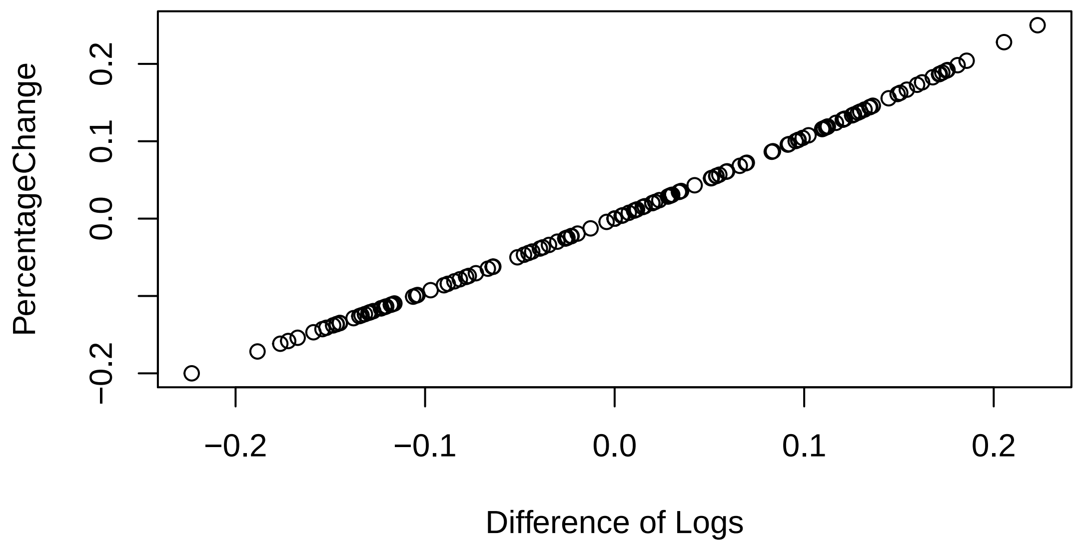
```r
plot(log(airpass),type='o',ylab='Log(Airline Passengers)')
```



Now the variation is similar at both high, low, and middel levels of the series.

(c):

```
percentage=na.omit((airpass-zlag(airpass))/zlag(airpass))
plot(x=diff(log(airpass))[-1],y=percentage[-1],ylab='PercentageChange',xlab='Difference of Logs')
```



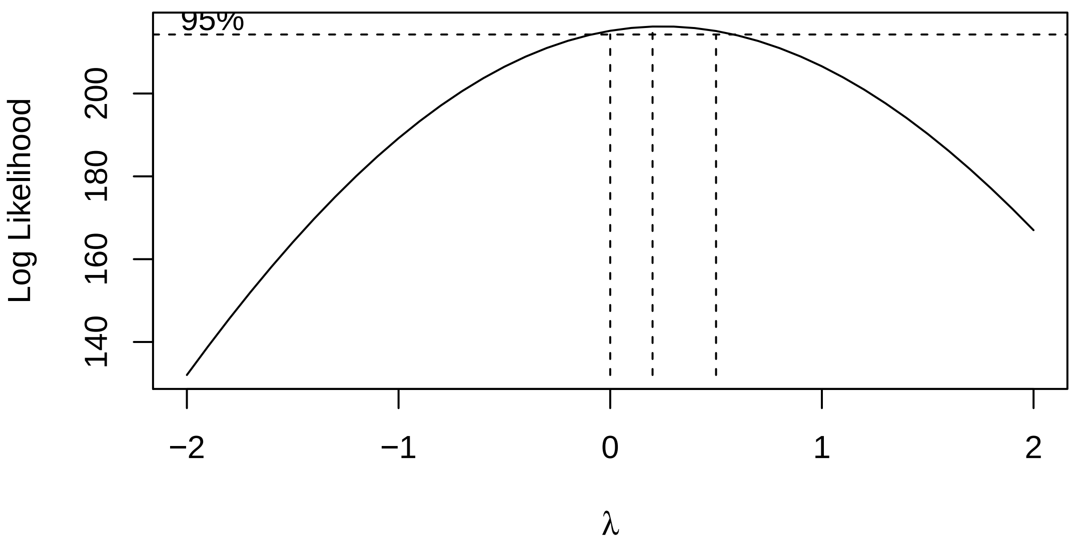
```r
cor(diff(log(airpass))[-1],percentage[-1])
```

```
## [1] 0.9986814
```

There is excellent agreement between the two transformed series in this class. The correlation coefficient in this plot is 0.999. Either transformation would be extremely helpful in modeling this series further.

5.14 (a)

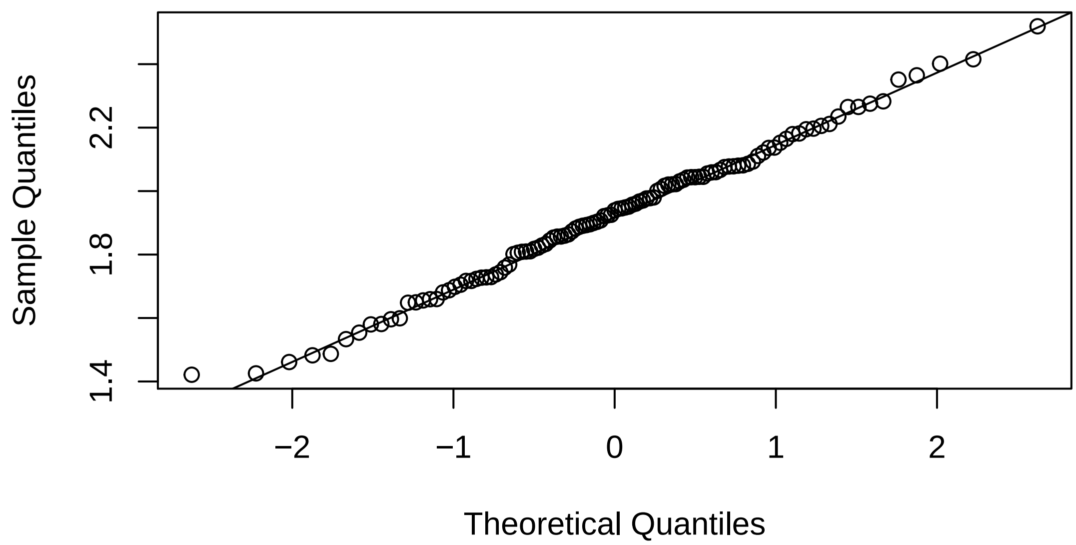
```r
data("larain")
BoxCox.ar(larain, method='ols')
```



The maximum likelihood value for lambda is about 0.26 but the 95% confidence interval includes the logarithm transformation(lambda=0) and the square root transformation(lambda=0.5). We choose lambda=0.25 or fourth root for the remaining section of this exercise.

(b)

```r
qqnorm((larain)^.25,main='')
qqline((larain)^.25)
```



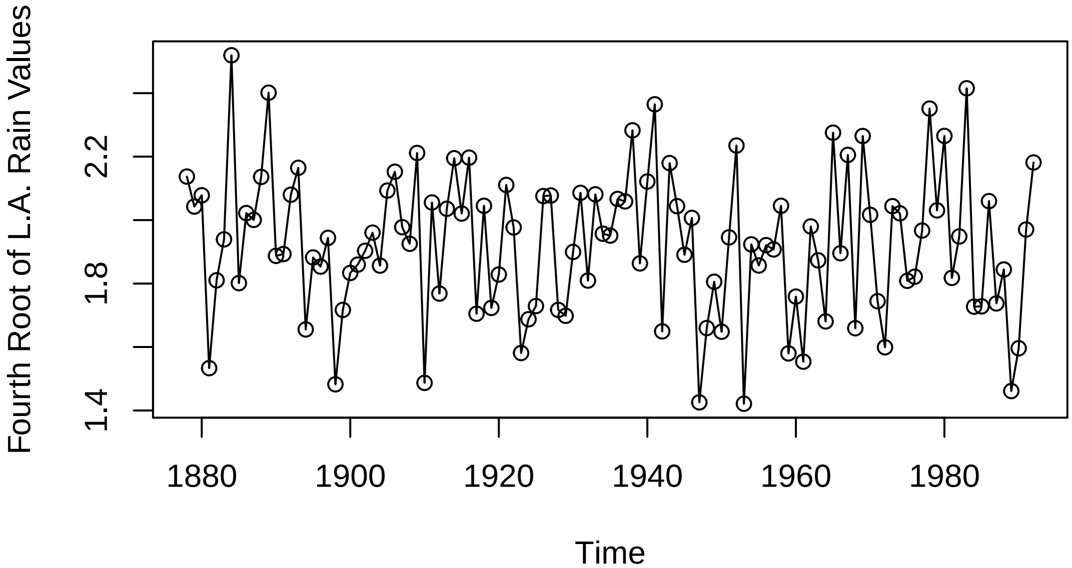

```r
shapiro.test((larain)^.25)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  (larain)^0.25
## W = 0.99408, p-value = 0.9096
```

The values transformed by the fourth root looks quite normal.

(c)

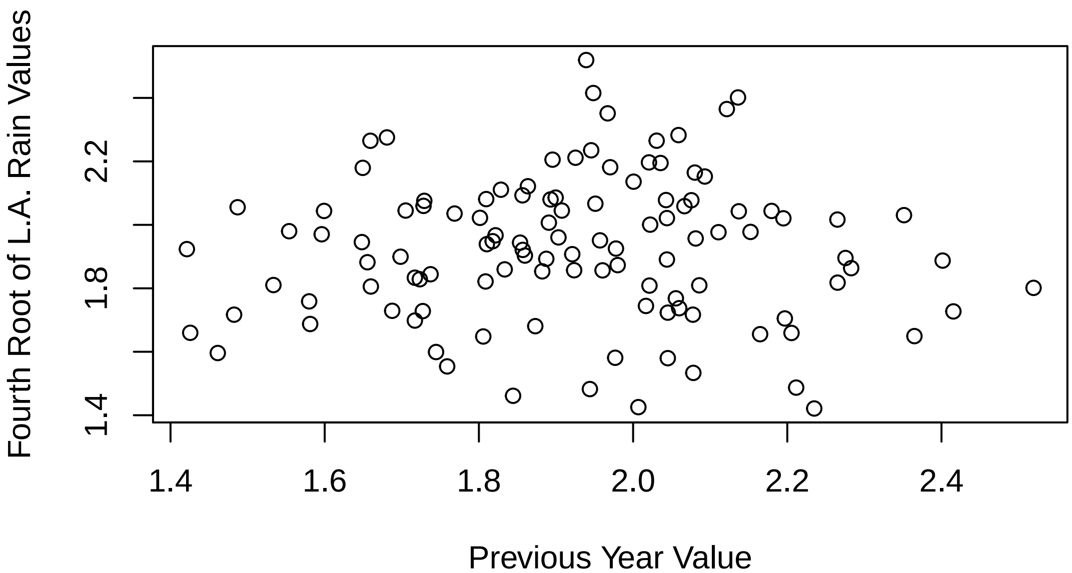
```r
plot(larain^.25,type='o',ylab='Fourth Root of L.A. Rain Values')
```



This transformed series could now be considered as normal white noise with a nonzero mean.

(d)

```
plot(y=(larain)^.25,x=zlag((larain)^.25),ylab='Fourth Root of L.A. Rain Values',
xlab='Previous Year Value')
```



The

lack of correlation or any other kind of dependency between year values is clear from this plot. Instantaneous transformations cannot induce correlation where none was present.

5 (i)

```
set.seed(1)
yt=arima.sim(n=100,list(ma=c(-0.7,-0.5,-0.6)))
print(acf(yt,plot=FALSE)$acf[4])
```

```
## [1] -0.09994342
```

(ii) As given in equation (4.25) and (6.1.11) in the textbook, we have:

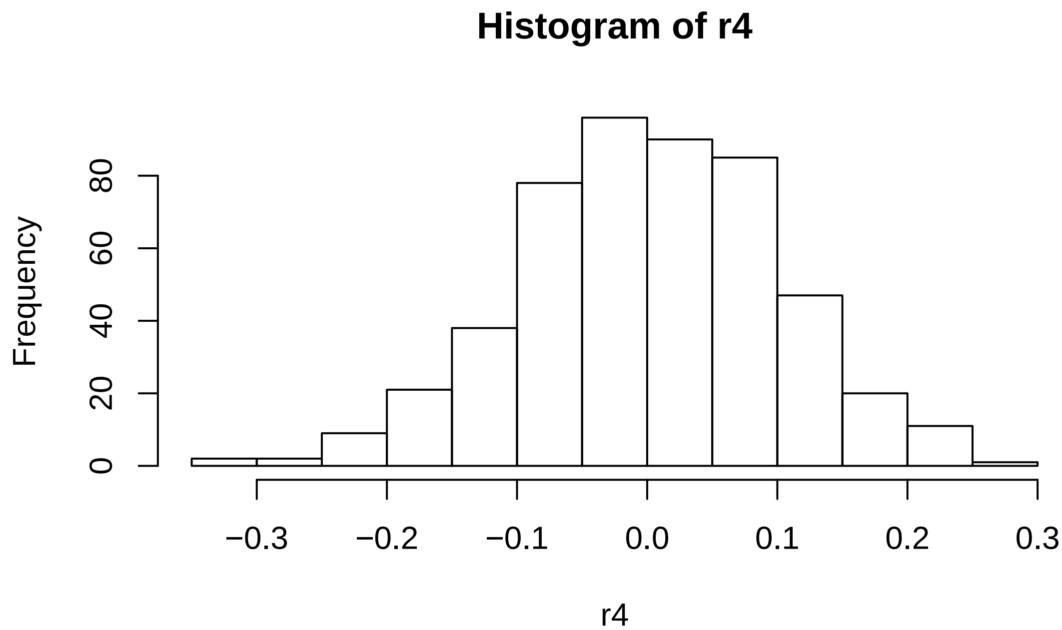
$$\rho_1 = \frac{-0.7 + 0.7 \times 0.5 + 0.5 \times 0.6}{1 + 0.7^2 + 0.5^2 + 0.6^2} = -0.02381$$
$$\rho_2 = \frac{-0.5 + 0.7 \times 0.6}{1 + 0.7^2 + 0.5^2 + 0.6^2} = -0.03810$$
$$\rho_3 = \frac{-0.6}{1 + 0.7^2 + 0.5^2 + 0.6^2} = -0.23810 \text{(or simply use the R code ARMAacf to calculate the theoretical ACF)}$$
$$\mathrm{Var}(r_4) = \frac{1}{n}[1 + 2\sum_{j=1}^{3} \rho_j^2] = \frac{1}{100}(1 + 2 \times -0.02381 + 2 \times -0.03810 + 2 \times -0.23810) = 0.01167$$

(iii)

```
r4=rep(0,500)
for (i in 1:500) {
xt=arima.sim(n=100,list(ma=c(-0.7,-0.5,-0.6)))
r4[i]=acf(xt,plot=FALSE)$acf[4]
}
print(var(r4))
```

```
## [1] 0.009895135
```

```
hist(r4)
```



(iv) use r_k in place of ρ_k in (6.1.11),we have:

$$\widehat{\text{Var}}(r_4) = \frac{1}{n}[1 + 2\sum_{j=1}^{3} r_j^2]$$

And the result is give by the following R code:

```
print(1/100*(1+2*sum(acf(yt,plot=FALSE)$acf[1:3]^2)))
```

```
## [1] 0.01219351
```

(v) For this problem we use the estimated variance of r_4 in the previous part, and since we know that r_4 is approximately normally distributed with mean ρ_4.Then for a 95% percent confidence interval, we have that :

$$\begin{aligned}
&\Phi^{-1}(0.025) \le \frac{r_4 - \rho_4}{\sqrt{\widehat{\text{Var}}(r_4)}} \le \Phi^{-1}(0.975) \\
\implies \quad & r_4 - \sqrt{\widehat{\text{Var}}(r_4)}\Phi^{-1}(0.975) \le \rho_4 \le r_4 - \sqrt{\widehat{\text{Var}}(r_4)}\Phi^{-1}(0.025) \\
\implies \quad & -0.1238423 \le \rho_4 \le -0.07604458
\end{aligned}$$

\textbf(iv) Since 0 does not fall into the confidence interval, the result is not consistent with the model.