HW4 Jin Kweon (3032235207)

Jin Kweon

3/1/2018

Problem 4

Data import

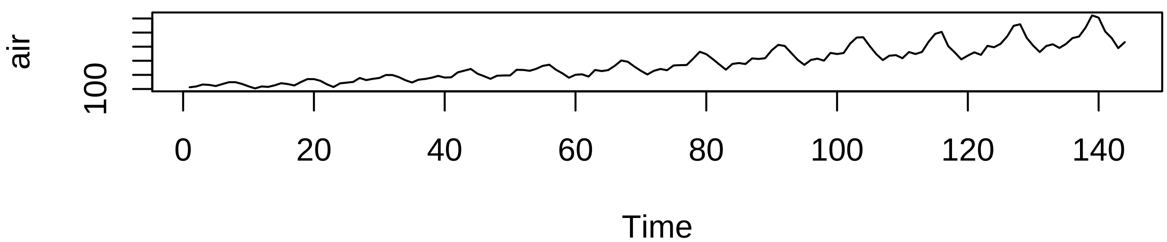
```
air <- read.table("airpass.txt", header = T)
air <- as.vector(air$airpass)
```

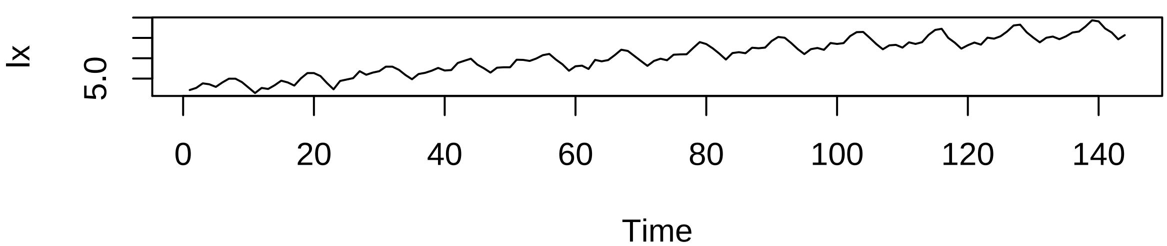
Part a

```
#plot.ts(AirPassengers) -> the dataset is already in R

#original
par(mfrow=c(2,1))
plot.ts(air)

#Logarithms of the series
lx <- log(air)
plot.ts(lx)
```



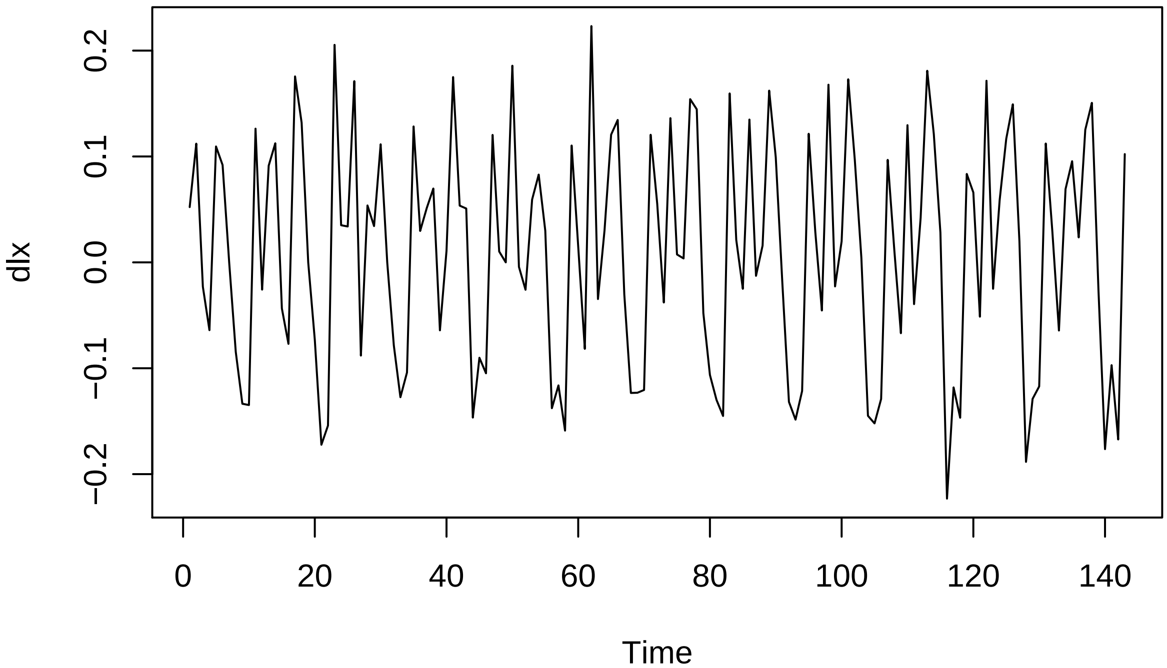


Comment:

I realized when I displayed the original model, the variance increases as the time goes. (and, that is one of the reason I can say that the model cannot be stationary) By taking logs (Box-cox power transformation), we are trying to stabilize the variance.

Part b

```
dlx <- diff(lx, 1, 1)
plot.ts(dlx)
```

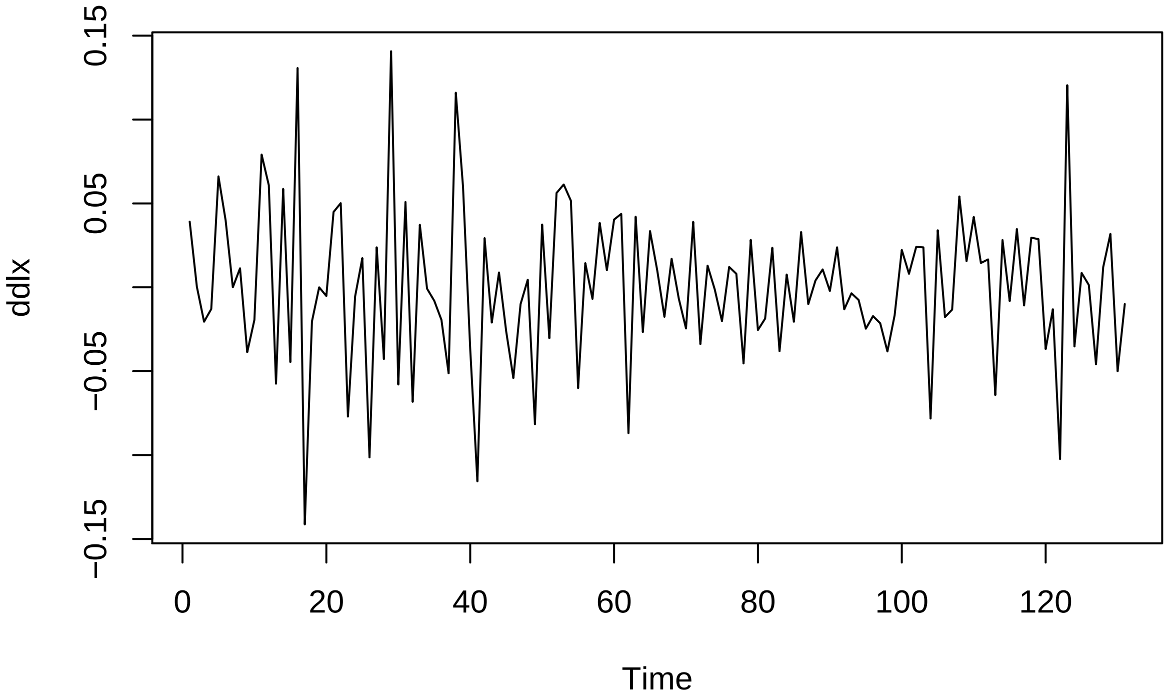


Comment:

Now, after I performed the first difference on the logged series, I can see the trend is removed. (now, the mean is more closed to constant, which is getting closer to stationarity) However, there is still seasonality I can find.

Part c

```
ddlx <- diff(dlx, 12, 1)
plot.ts(ddlx)
```

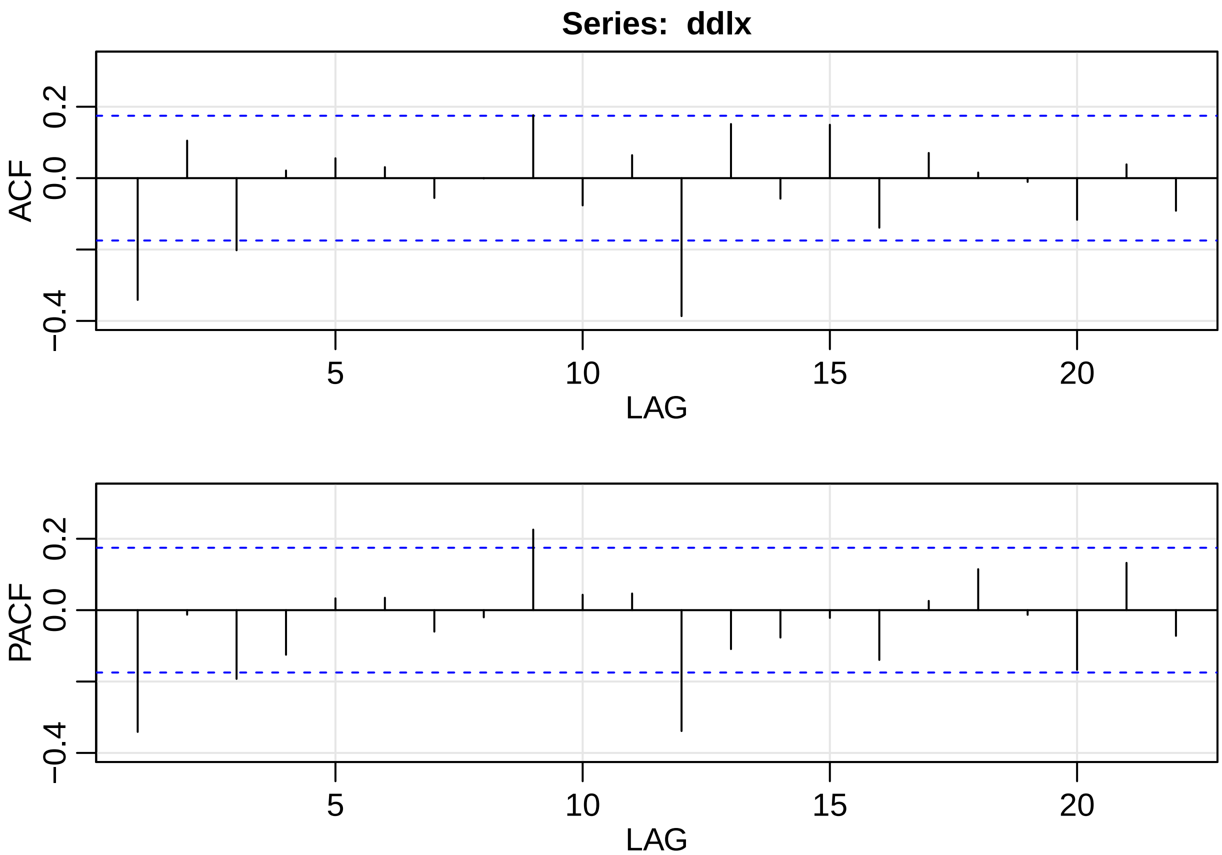


Comment:

I did twelfth-order differencing to remove the seasonality. (autocovariance is dependent on time only through absolute value of lag, and mean is constant)

Part d & e

```r
acf2(ddlx)
```

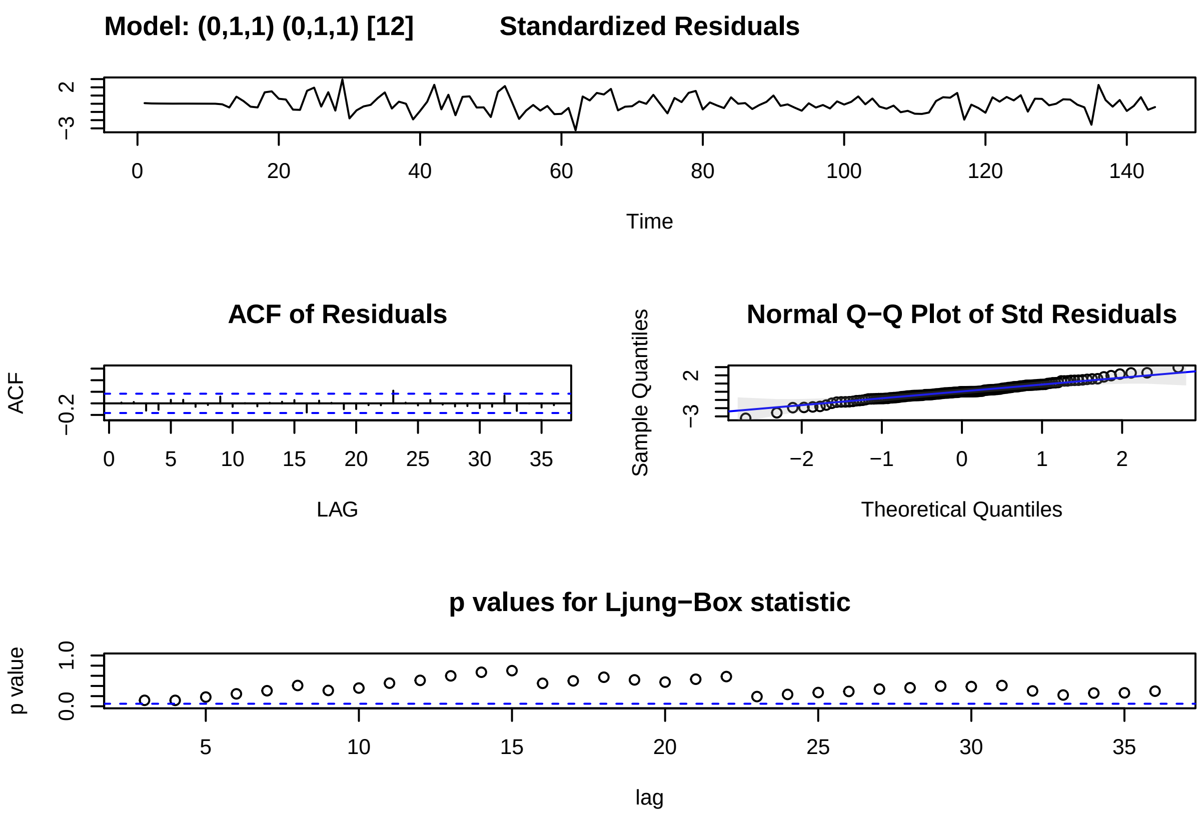



```
##         ACF  PACF
##  [1,] -0.34 -0.34
##  [2,]  0.11 -0.01
##  [3,] -0.20 -0.19
##  [4,]  0.02 -0.13
##  [5,]  0.06  0.03
##  [6,]  0.03  0.03
##  [7,] -0.06 -0.06
##  [8,]  0.00 -0.02
##  [9,]  0.18  0.23
## [10,] -0.08  0.04
## [11,]  0.06  0.05
## [12,] -0.39 -0.34
## [13,]  0.15 -0.11
## [14,] -0.06 -0.08
## [15,]  0.15 -0.02
## [16,] -0.14 -0.14
## [17,]  0.07  0.03
## [18,]  0.02  0.11
## [19,] -0.01 -0.01
## [20,] -0.12 -0.17
## [21,]  0.04  0.13
## [22,] -0.09 -0.07
```

```
sarima(lx, 0, 1, 1, 0, 1, 1, 12)
```

```
## initial  value -3.086228
## iter   2 value -3.267980
```

```
## iter   3 value -3.279950
## iter   4 value -3.285996
## iter   5 value -3.289332
## iter   6 value -3.289665
## iter   7 value -3.289672
## iter   8 value -3.289676
## iter   8 value -3.289676
## iter   8 value -3.289676
## final  value -3.289676
## converged
## initial  value -3.286464
## iter   2 value -3.286855
## iter   3 value -3.286872
## iter   4 value -3.286874
## iter   4 value -3.286874
## iter   4 value -3.286874
## final  value -3.286874
## converged
```




```
## $fit
##
## Call:
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,
##     Q), period = S), include.mean = !no.constant, optim.control = list(trace = trc,
##     REPORT = 1, reltol = tol))
##
## Coefficients:
##           ma1     sma1
##       -0.4018  -0.5569
```

```
## s.e.   0.0896   0.0731
##
## sigma^2 estimated as 0.001348:  log likelihood = 244.7,  aic = -483.4
##
## $degrees_of_freedom
## [1] 129
##
## $ttable
##      Estimate     SE t.value p.value
## ma1   -0.4018 0.0896 -4.4825       0
## sma1  -0.5569 0.0731 -7.6190       0
##
## $AIC
## [1] -5.58133
##
## $AICc
## [1] -5.56625
##
## $BIC
## [1] -6.540082
```

Comment:

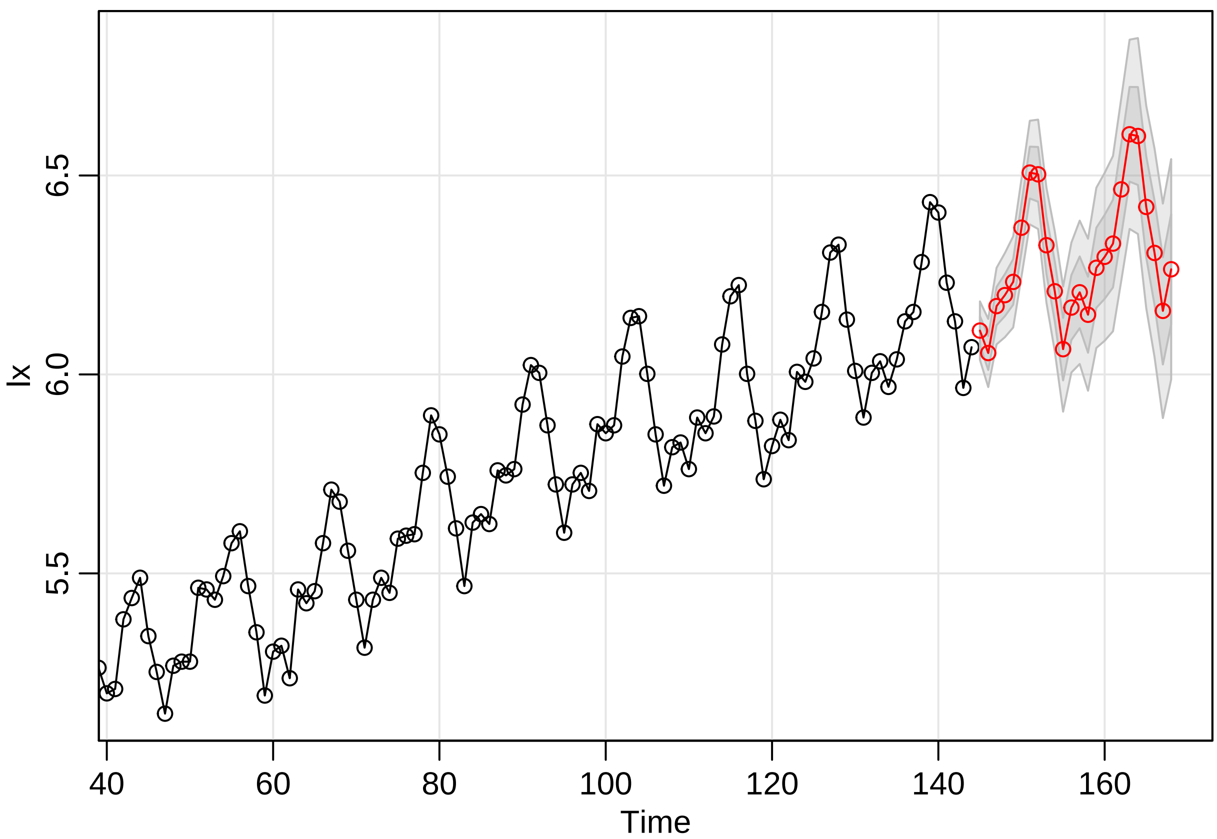
The model seems pretty good. (based on different model diagnostic plots and ttable p-values)

ACF of residuals are mostly lying within $\pm\frac{1.96}{\sqrt{n}}$, meaning that residuals are uncorrelated.

And, normality assumption for standardized residual is pretty reasonable, as it is closed to a straight line.

Part f

```r
sarima.for(lx, 24, 0, 1, 1, 0, 1, 1, 12)
```




```
## $pred
## Time Series:
## Start = 145
## End = 168
## Frequency = 1
##  [1] 6.110186 6.053775 6.171715 6.199300 6.232556 6.368779 6.507294
##  [8] 6.502906 6.324698 6.209008 6.063487 6.168025 6.206435 6.150025
## [15] 6.267964 6.295550 6.328805 6.465028 6.603543 6.599156 6.420947
## [22] 6.305257 6.159737 6.264274
##
## $se
## Time Series:
## Start = 145
## End = 168
## Frequency = 1
##  [1] 0.03671562 0.04278291 0.04809072 0.05286830 0.05724856 0.06131670
##  [7] 0.06513124 0.06873441 0.07215787 0.07542612 0.07855851 0.08157070
## [13] 0.09008475 0.09549708 0.10061869 0.10549195 0.11014981 0.11461854
## [19] 0.11891946 0.12307018 0.12708540 0.13097758 0.13475740 0.13843405
```

Comment:

Based on the approximations, in the next two years, the number of passengers will slightly go up.