

# MIDS W271-4 Homework 7

*Subhashini R., Lei Yang, Ron Cordell*

*March 22, 2016*

## Question 1:

### 1.1 Load hw07\_series1.csv

```
df <- read.csv('hw07_series1.csv')
```

### 1.2 Describe the basic structure of the data and provide summary statistics of the series

```
# take a look at the data structure  
str(df)
```

```
## 'data.frame':    74 obs. of  1 variable:  
## $ X10.01: num  10.07 10.32 9.75 10.33 10.13 ...
```

```
# take a look at the first part of the data  
head(df,10)
```

```
##      X10.01  
## 1    10.07  
## 2    10.32  
## 3     9.75  
## 4    10.33  
## 5    10.13  
## 6    10.36  
## 7    10.32  
## 8    10.13  
## 9    10.16  
## 10   10.58
```

```
# get a summary of the data  
summary(df)
```

```
##      X10.01  
## Min.   : 9.75  
## 1st Qu.:10.48  
## Median :10.82  
## Mean   :10.82  
## 3rd Qu.:11.06  
## Max.   :11.94
```

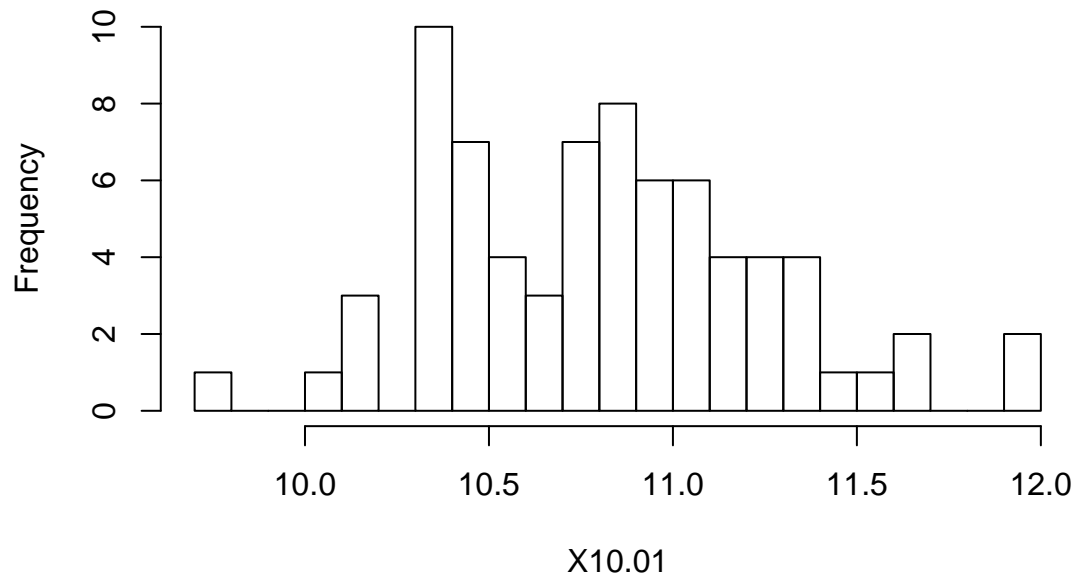
The data consists of 74 observations of a single variable, X10.01. There is no information on the sampling rate or if the samples fit into an overall time structure such as years, quarters, months or something else (frequency). Therefore we will create a time series object from it without imposing those parameters.

```
ts1 <- ts(df, start=1, end=74, frequency=1)
```

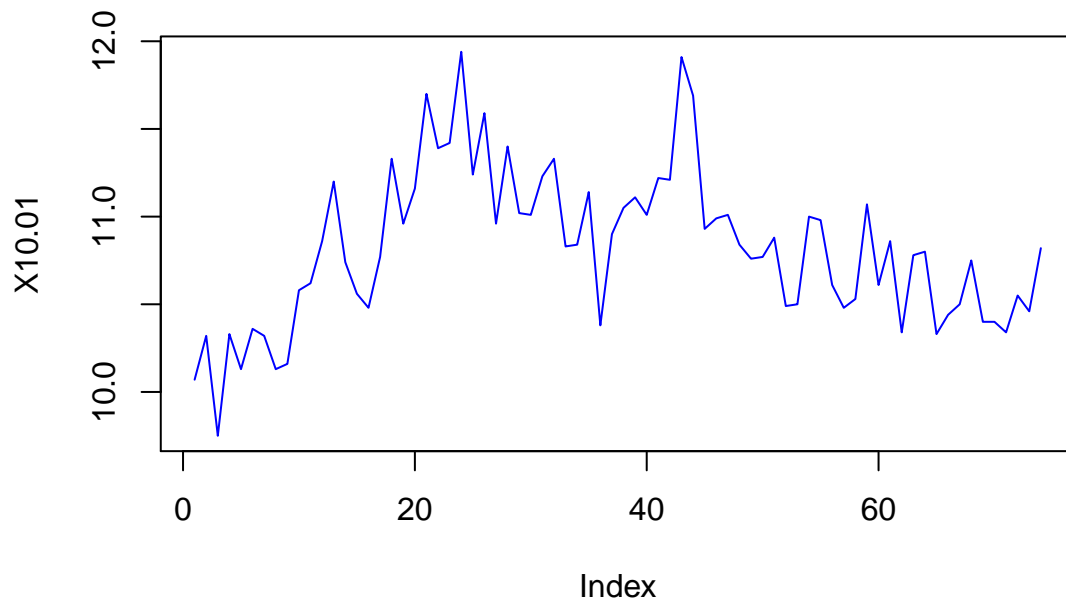
1.3 Plot histogram and time-series plot of the series. Describe the patterns exhibited in histogram and time-series plot. For time series analysis, is it sufficient to use only histogram to describe a series?

```
par(mfrow=c(2,1))  
hist(ts1, main='Histogram of Time Series', xlab='X10.01', breaks=20)  
plot.ts(ts1, col='blue',  
        xlab='Index',  
        ylab='X10.01',  
        main='Time Series Plot of X10.01 Series')
```

**Histogram of Time Series**

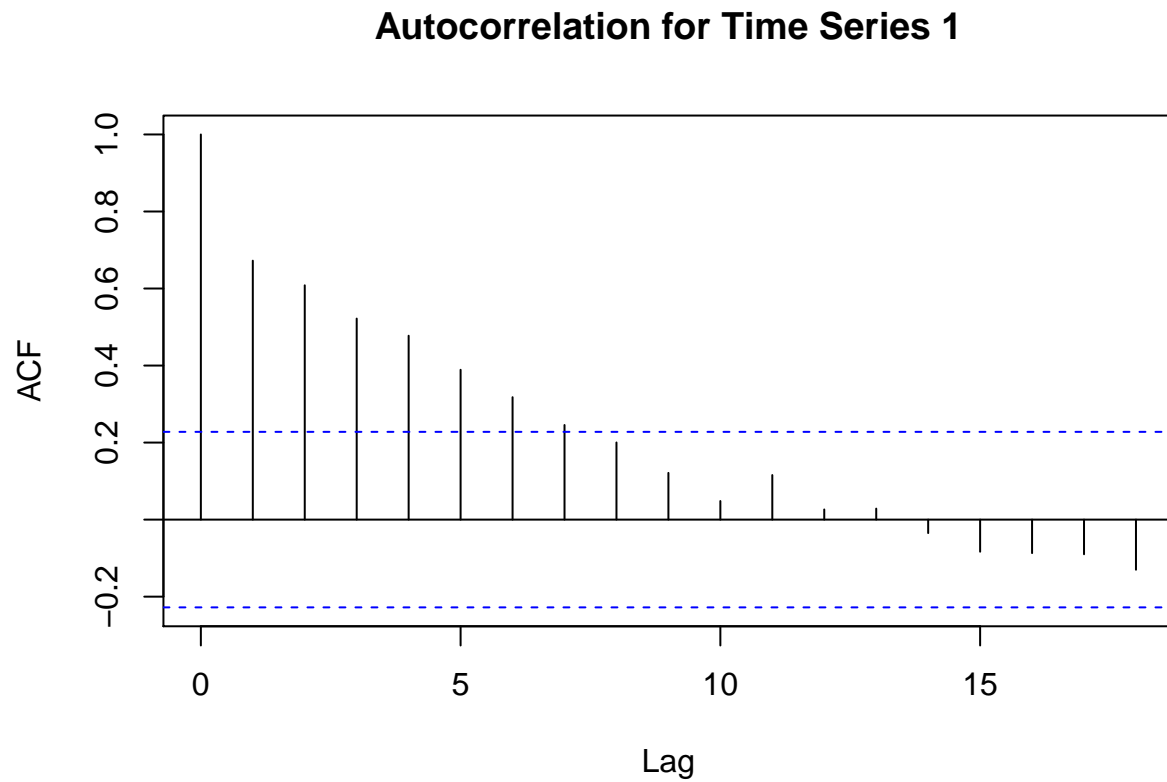


**Time Series Plot of X10.01 Series**



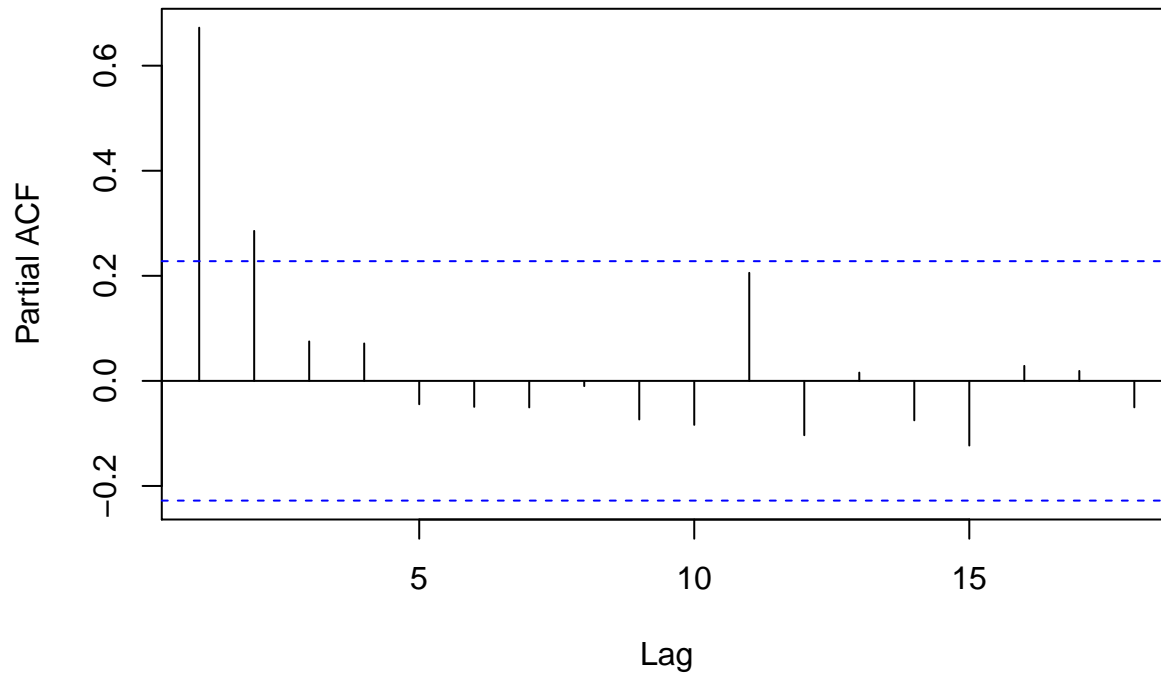
1.4 Plot the ACF and PACF of the series. Describe the patterns exhibited in the ACF and PACF.

```
acf(ts1, xlab='Lag', main='Autocorrelation for Time Series 1')
```



```
pacf(ts1, xlab='Lag', main='Partial Autocorrelation for Time Series 1')
```

## Partial Autocorrelation for Time Series 1



1.5 Estimate the series using the `ar()` function.

```
ts1.ar <- ar(ts1, model='mle')
```

1.6 Report the estimated AR parameters, the order of the model, and standard errors.

```
# AR model estimated parameters
ts1.ar$ar
```

```
## [1] 0.4803726 0.2854828
```

```
# AR model order
ts1.ar$order
```

```
## [1] 2
```

```
# AR model standard errors
ts1.ar$asy.var.coef
```

```
##           [,1]      [,2]
## [1,] 0.012936613 -0.008697333
## [2,] -0.008697333 0.012936613
```

## Question 2:

2.1 Simulate a time series of length 100 for the following model. Name the series `x`.

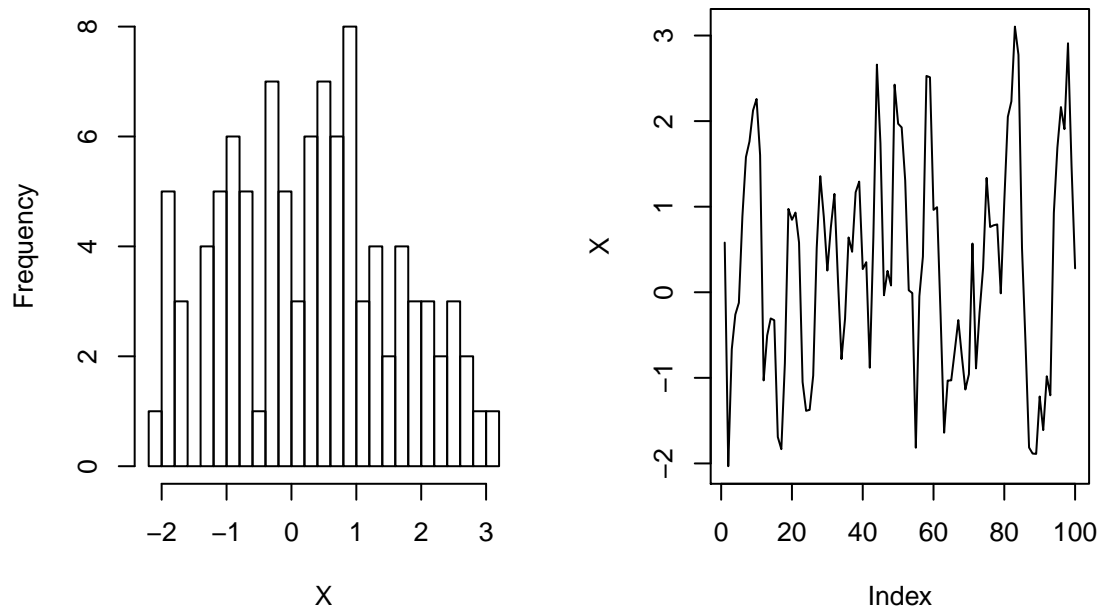
$$x_t = \frac{5}{6}x_{t-1} - \frac{1}{6}x_{t-2} + \omega_t$$

```
set.seed(1)
x <- arima.sim(n = 100, list(ar = c(5/6, -1/6), ma=0))
```

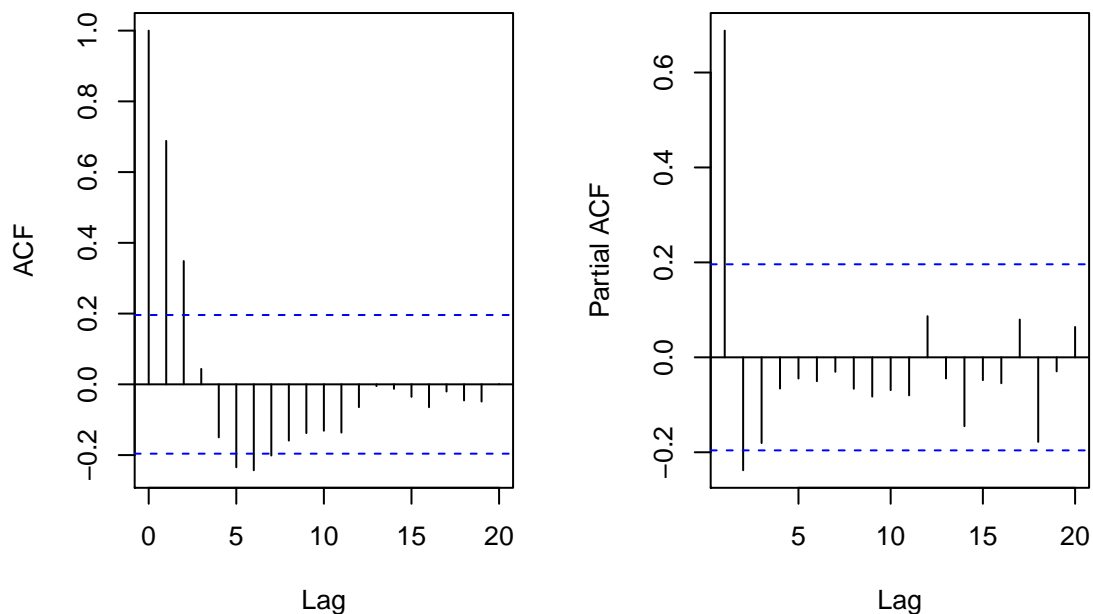
**2.2** Plot the correlogram and partial correlogram for the simulated series. Comments on the plots.

```
par(mfrow=c(2,2))
hist(x, main='Histogram of AR(2) Simulated Series',
     xlab='X', breaks=30)
plot.ts(x,
        main='Time Series Plot of AR(2) Simulated Series',
        xlab='Index', ylab='X')
acf(x, main='Correlogram of AR(2) Simulated Series',
    xlab='Lag')
pacf(x,
     xlab='Lag',
     main='Partial Correlogram of AR(2) Simulated Series')
```

**Histogram of AR(2) Simulated Series Time Series Plot of AR(2) Simulated Series**



**Correlogram of AR(2) Simulated Series Partial Correlogram of AR(2) Simulated Series**



The histogram of the simulated AR(2) series shows a somewhat symmetrical distribution about a mean of 0 but it doesn't look normally distributed. The time series plot shows downward and upward trends over some time indices but no persistent trend is apparent. The correlogram of the simulated series shows an quickly decaying but oscillating trend to 0. The partial correlogram of the simulated series shows a rapid dropoff after the first lag.

**2.3 Estimate an AR model for this simulated series. Report the estimated AR parameters, standard errors, and the order of the AR model.**

```
# fit an AR model to the X series
x.ar <- ar(x, method='mle')
```

```
# estimated coefficients
x.ar$ar
```

```
## [1] 0.81405014 -0.08889983 -0.19629863
```

```
# show the model order
x.ar$order
```

```
## [1] 3
```

```
# show the parameter estimate standard errors
x.ar$asy.var.coef
```

```
##           [,1]      [,2]      [,3]
## [1,] 0.009553607 -0.008139448 0.002272339
## [2,] -0.008139448 0.015947746 -0.008139448
## [3,] 0.002272339 -0.008139448 0.009553607
```

```
# show the model AIC output
x.ar$aic
```

```
##           0           1           2           3           4           5           6
## 67.686495 6.093096 1.504420 0.000000 1.591450 3.305639 4.960726
##           7           8           9          10          11          12
## 6.826557 8.405156 9.892481 11.624438 12.980320 14.008494
```

The model is estimated as an AR(3) model. However, examining the AIC output of the `ar()` command shows that a more parsimonious AR(2) model may be available with a very slightly larger AIC.

**2.4 Construct a 95% confidence intervals for the parameter estimates of the estimated model. Do the “true” model parameters fall within the confidence intervals? Explain the 95% confidence intervals in this context.**

```
# Confidence Interval of the AR parameters:
x.ar$ar + matrix(c(2,-2, 2, -2, 2, -2),
                 nrow = 3, ncol = 2, byrow = TRUE)*
  sqrt(diag(x.ar$asy.var.coef))
```

```
##           [,1]      [,2]
## [1,] 1.0095352505 0.6185650
## [2,] 0.1636689405 -0.3414686
## [3,] -0.0008135205 -0.3917837
```

```
# Actual model parameters:
```

```
5/6
```

```
## [1] 0.8333333
```



```
-1/6
```

```
## [1] -0.1666667
```

```
0
```

```
## [1] 0
```

The first two actual AR process parameters are contained in the confidence intervals of their estimated counterparts but the third parameter is not.

## 2.5 Is the estimated model stationary or non-stationary?

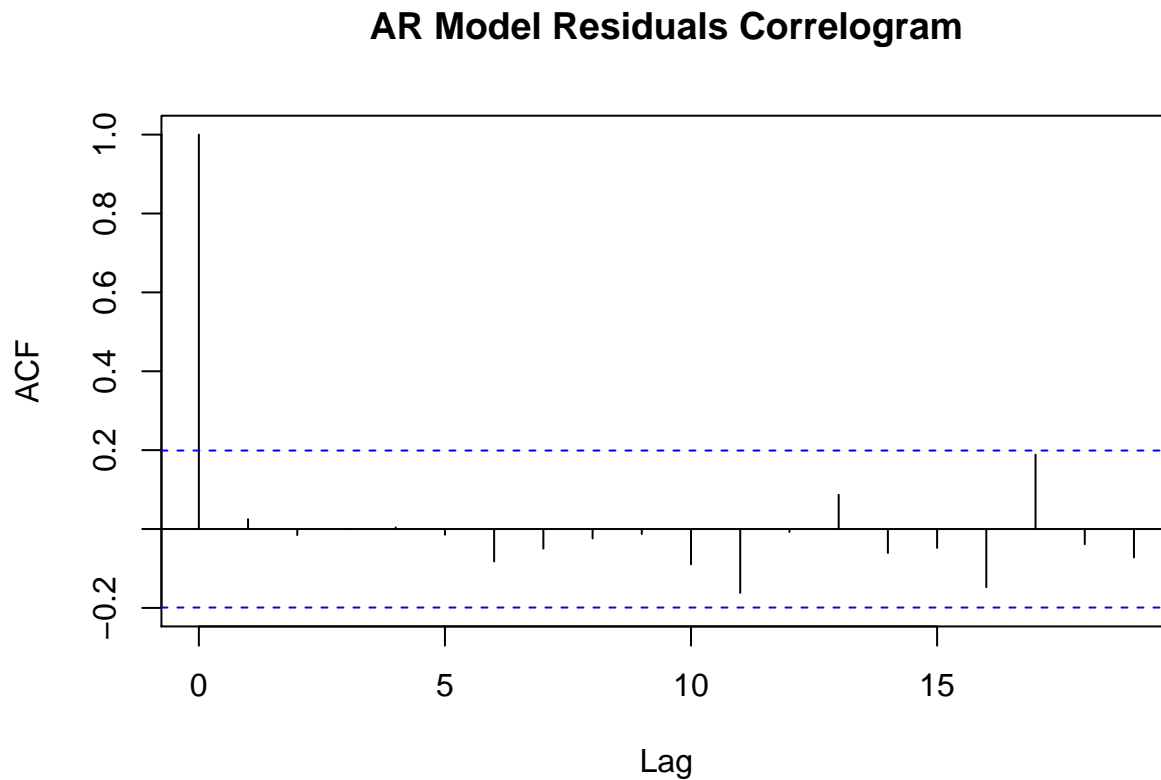
```
# Find the absolute values of the roots of the lagged polynomial  
abs(polyroot(c(1, -0.81405014, 0.08889983, 0.19629863)))
```

```
## [1] 1.375108 1.375108 2.694070
```

The absolute values of the roots of the lagged polynomial are all larger than 1 therefore the estimated model is stationary.

## 2.6 Plot the correlogram of the residuals of the estimated model. Comment on the plot.

```
acf(x.ar$resid, na.action=na.omit, main='AR Model Residuals Correlogram')
```



The correlogram shows no autocorrelation in the residuals.