Trend Analysis

Nik Bear Brown

In this lesson we'll learn the how to implement Trend Analysis in R.

# Additional packages needed

To run the code you may need additional packages.

* If necessary install the followings packages.

install.packages("tseries");

library(tseries)

# Data

We will be generating simulated data for this lesson.

# Trend Analysis

Trend Analysis is the practice of collecting information and attempting to spot a pattern, or trend, in the information. Typically this involves analyzing the variance for a change over time. The null hypothesis: is that there is no trend. Many techniques can be used to identify trends, we'll use an ARMA model again.

# Dickey-Fuller Test

The Dickey-Fuller Test is a test for the stationarity of a time series.

The [Dickey-Fuller test](https://en.wikipedia.org/wiki/Dickey%E2%80%93Fuller_test) tests whether a unit root is present in an autoregressive model. simple AR(1) model is

where is the variable of interest, is the time index, is a coefficient, and is the error term. A unit root is present if . The model would be non-stationary in this case.

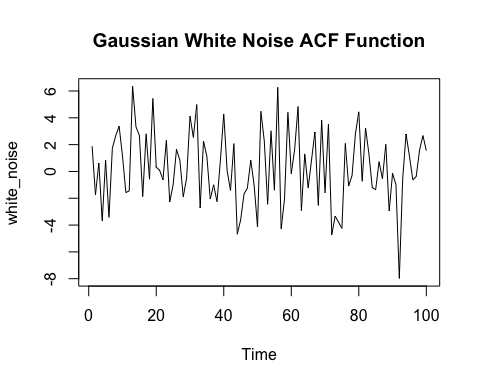
The regression model can be written as

where is the first difference operator.

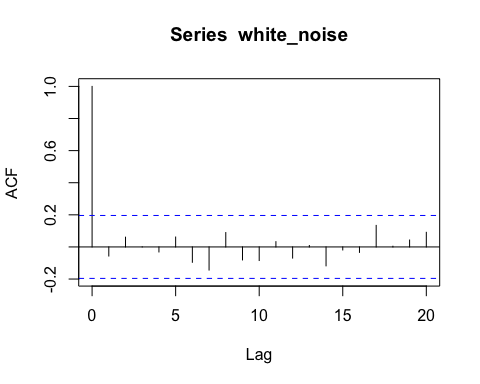
The Dickey-Fuller Test uses a specific distribution simply known as the Dickey-Fuller table to assess whether is signficant.

# Trend Analysis in R

#----------- White noise --------  
set.seed(3333)  
white\_noise <- rnorm(100, mean = 0, sd = 3.0)  
plot(white\_noise,type='l',xlab='Time',main='Gaussian White Noise ACF Function')



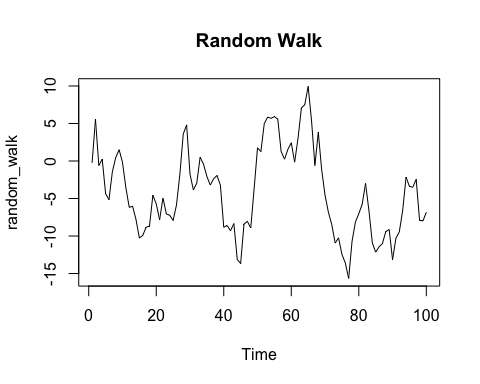
## Fitting White noise Time series   
# plot a correlogram  
acf(white\_noise)



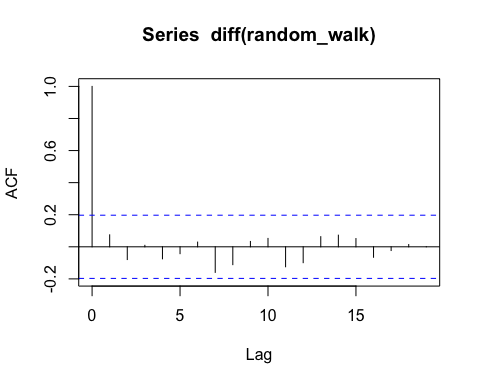
var(white\_noise)

## [1] 7.576885

#------------- Random Walk -------------  
#To simulate a random walk in R, we essentially need a cumulative sum of a white noise random series.   
set.seed(333)  
random\_walk <- cumsum(rnorm(100, mean = 0, sd = 3.0))  
plot(random\_walk,type='l',xlab='Time',main='Random Walk')



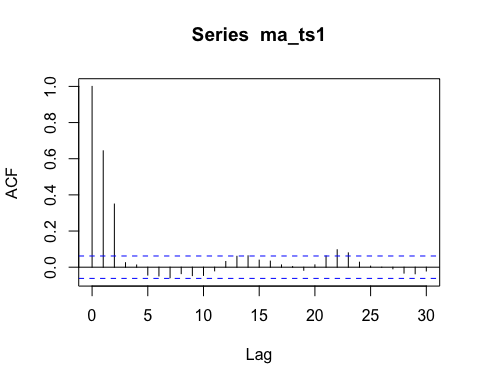
# Fitting a Random Walk   
#A good way to see if a time series follows a random walk is to compute the successive differences between terms.  
acf(diff(random\_walk))



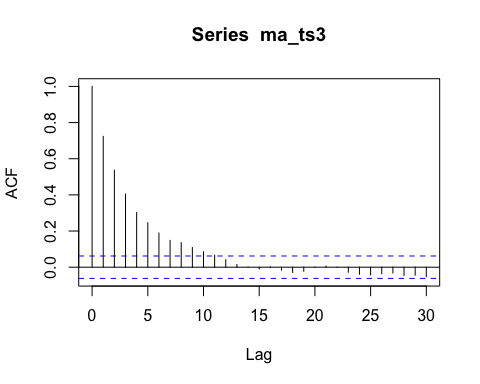
#----------------------------- ARMA model ----------------------  
  
 # Moving Average Model  
  
set.seed(555)  
#ACF function with coefficients 0.84 and 0.62  
ma\_ts1 <- arima.sim(model = list(ma = c(0.84, 0.62), sd = 1.2), n = 1000)  
head(ma\_ts1, n = 8)

## [1] 0.59291917 2.51535102 0.03864703 0.56145710 -0.51638224 1.78195701  
## [7] 1.08566382 1.49532881

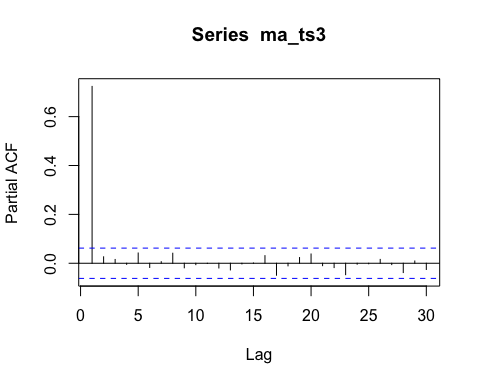
acf(ma\_ts1)



# Autoregressive model  
set.seed(5555)  
ma\_ts3 <- arima.sim(model = list(ar = c(0.74), sd = 1.2), n = 1000)  
acf(ma\_ts3)



pacf(ma\_ts3)



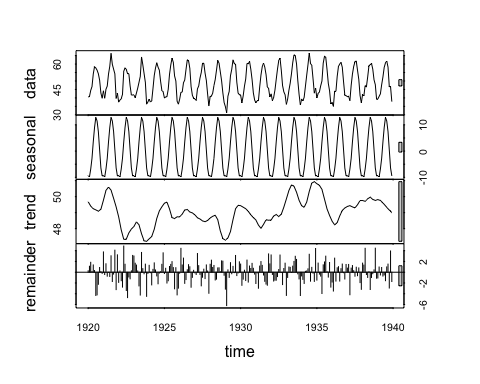
#--------------Dickey-Fuller for stationarity -----------------------  
adf.test(random\_walk, alternative = "stationary")

##   
## Augmented Dickey-Fuller Test  
##   
## data: random\_walk  
## Dickey-Fuller = -2.7267, Lag order = 4, p-value = 0.2756  
## alternative hypothesis: stationary

#------------ Another unit root test : Philips-Perron test -------  
PP.test(random\_walk)

##   
## Phillips-Perron Unit Root Test  
##   
## data: random\_walk  
## Dickey-Fuller = -2.8444, Truncation lag parameter = 3, p-value =  
## 0.2268

# ------------ Seasonal Trend Decomposition in R --------  
  
#The Seasonal Trend Decomposition using Loess (STL) is an algorithm that was developed   
#to help to divide up a time series into three components namely: the trend, seasonality and remainder.  
nottem.stl = stl(nottem, s.window="periodic")  
plot(nottem.stl)



# Resources

* The heat is on.. or is it? Trend Analysis of Toronto Climate Data via @rbloggers](<http://www.r-bloggers.com/the-heat-is-on-or-is-it-trend-analysis-of-toronto-climate-data/>)
* [Trend Analysis - ETH](http://www.iac.ethz.ch/edu/courses/master/electives/acwd/Trend.pdf)
* [Trend Analysis Using R - ResearchGate](http://www.researchgate.net/publication/275640899_Trend_Analysis_Using_R)

# References

The data, R code and lessons are based upon:

1. Time Series Analysis :

Data Source: <http://www.geophysics.geol.uoa.gr/catalog/catgr_20002008.epi>

Code References :

Book : Mastering Predictive Analytic with R  
Author: Rui Miguel Forte  
<https://www.safaribooksonline.com/library/view/mastering-predictive-analytics/9781783982806/>

Chapter 9: Time series Analysis

<http://www.statoek.wiso.uni-goettingen.de/veranstaltungen/zeitreihen/sommer03/ts_r_intro.pdf>

<http://www.stat.pitt.edu/stoffer/tsa3/R_toot.htm>

<http://www.statoek.wiso.uni-goettingen.de/veranstaltungen/zeitreihen/sommer03/ts_r_intro.pdf>

1. Trend Analysis

Code References :

Book : Mastering Predictive Analytic with R  
Author: Rui Miguel Forte  
<https://www.safaribooksonline.com/library/view/mastering-predictive-analytics/9781783982806/>

<http://www.r-bloggers.com/seasonal-trend-decomposition-in-r/>

1. Seasonal Models

Code references :

Book: Time Series Analysis and Its Applications  
Author: Robert H. Shumway . David S. Stoffer  
Link: <http://www.springer.com/us/book/9781441978646#otherversion=9781461427599>

<http://a-little-book-of-r-for-time-series.readthedocs.org/en/latest/src/timeseries.html>

<https://onlinecourses.science.psu.edu/stat510/?q=node/47>

<https://rpubs.com/ryankelly/tsa5>

<https://onlinecourses.science.psu.edu/stat510/node/68>

Data Reference : <https://github.com/RMDK/TimeSeriesAnalysis/blob/master/colorado_river.csv>

1. Spectral Analysis

Code References:  
Book:  
Modern Applied Statistics with S Fourth edition  
Author: W. N. Venables and B. D. Ripley  
Link: Modern Applied Statistics with S Fourth edition

<http://www.maths.adelaide.edu.au/patty.solomon/TS2004/tsprac3_2004.pdf>