# stat_626_assignment02

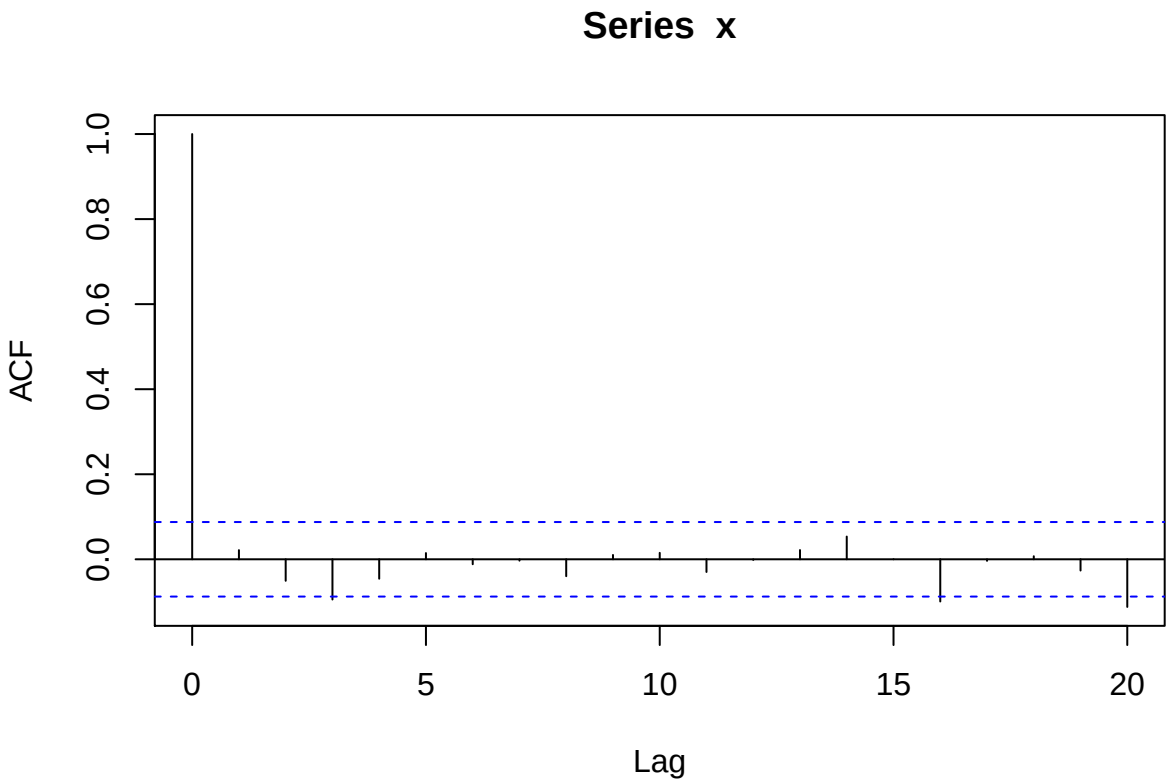
*Philip Anderson; panders2@tamu.edu*

*6/1/2018*

## Question 1.15

### 1.15 A

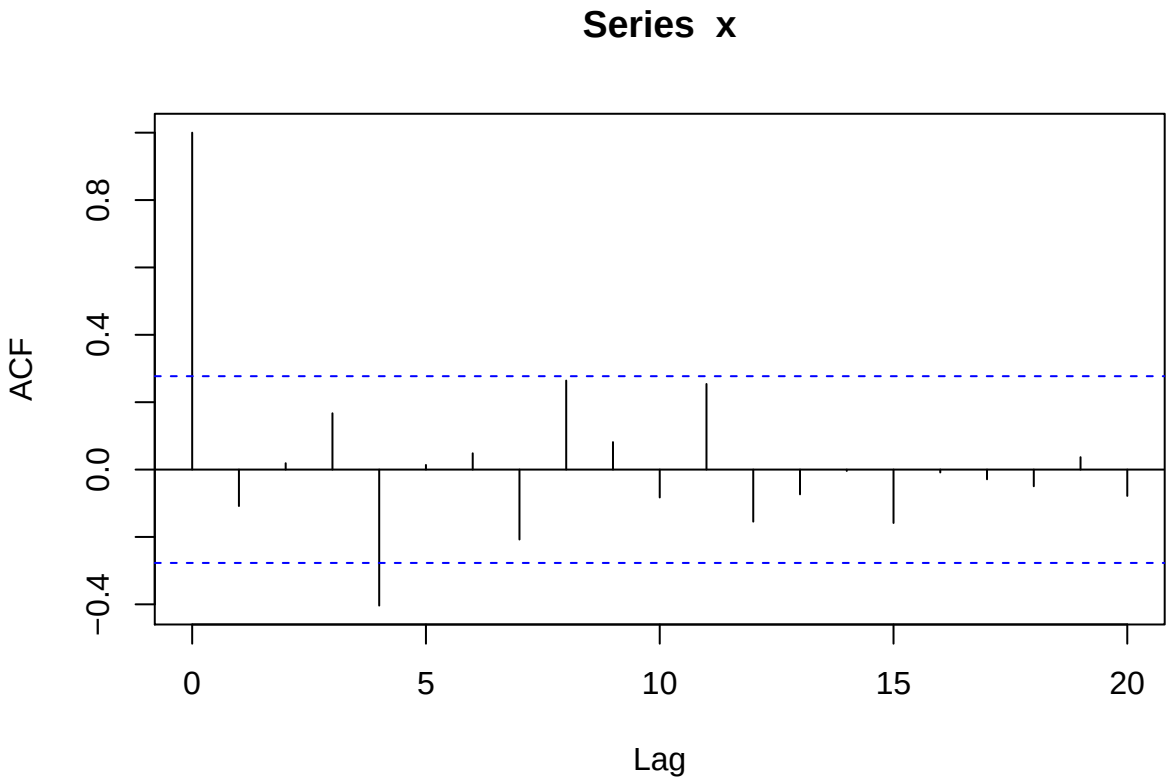
```
# 500 random draws from a white noise process
x <- rnorm(500, 0, 1)
# produce autocorrelation function
acf(x, lag.max=20)
```




The theoretical ACF for white noise holds that the autocorrelation will be 0 for any lag $h$ greater than 0. We can see from the autocorrelation function plotted above that there is no clear trend for the lags. Lag 23 is significant, but we would expect to see around 1 of every 20 lags produce a significant value by chance alone.

### 1.15 B

```
x <- rnorm(50, 0, 1)
acf(x, lag.max=20)
```



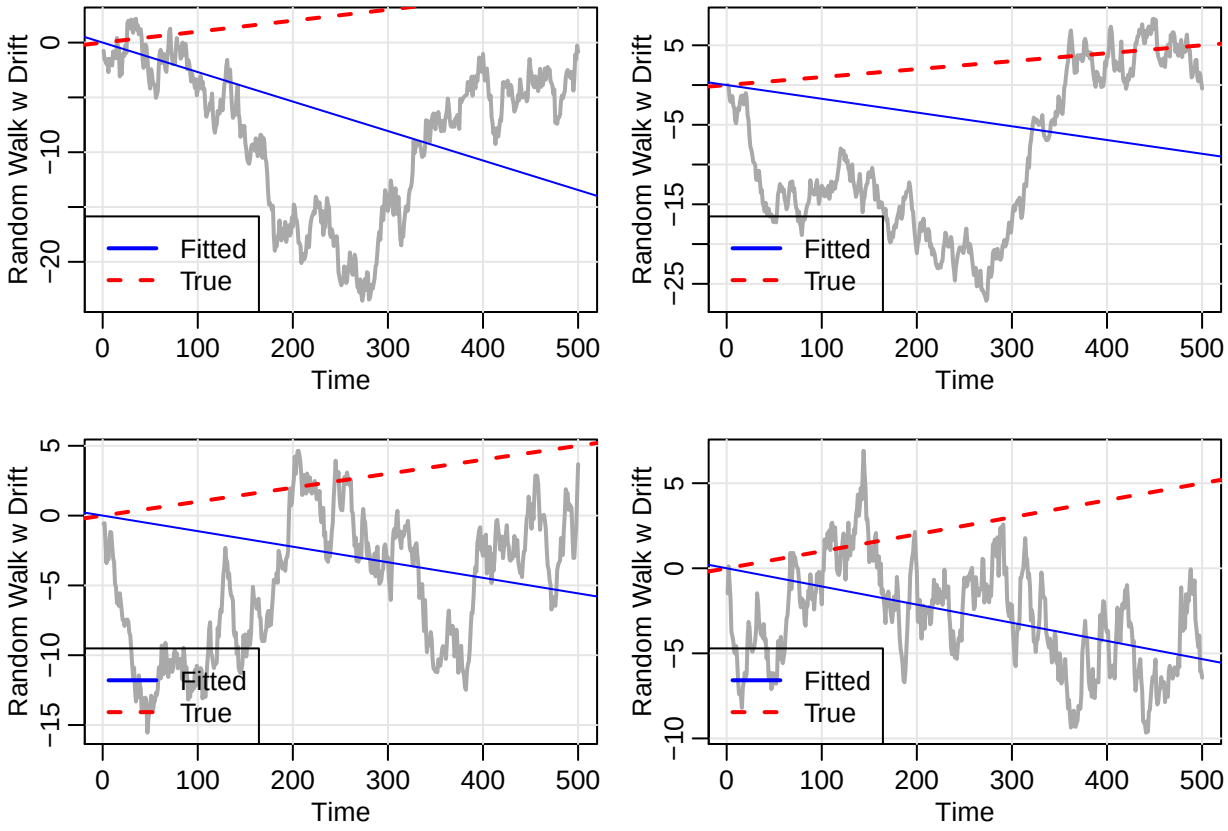

Drawing only 50 samples has produced similar results as far as patterns go, but has produced wider bands, likely due to an increase in variance from the reduced sample size.

## Question 2.3

### 2.3.A

```
# graphic set up
par(mfrow=c(2,2), mar=c(2.5,2.5,0,0)+.5, mgp=c(1.6,.6,0))

for (i in 1:4){
  set.seed(i+8)
  # generate random walk with drift
  x <- ts(cumsum(rnorm(500,.01,1)))
  # fit no-intercept regression
  regx = lm(x~0+time(x), na.action=NULL)
  # create time series plot object
  tsplot(x, ylab='Random Walk w Drift', col='darkgray', lwd=2)
  # true mean function
  abline(a=0, b=.01, col="red", lty=2, lwd=2)
  # regression plot
  abline(regx, col="blue")
  legend("bottomleft", c("Fitted", "True"), lty=c(1,2), lwd=c(2,2), col=c("blue","red"))
}
```

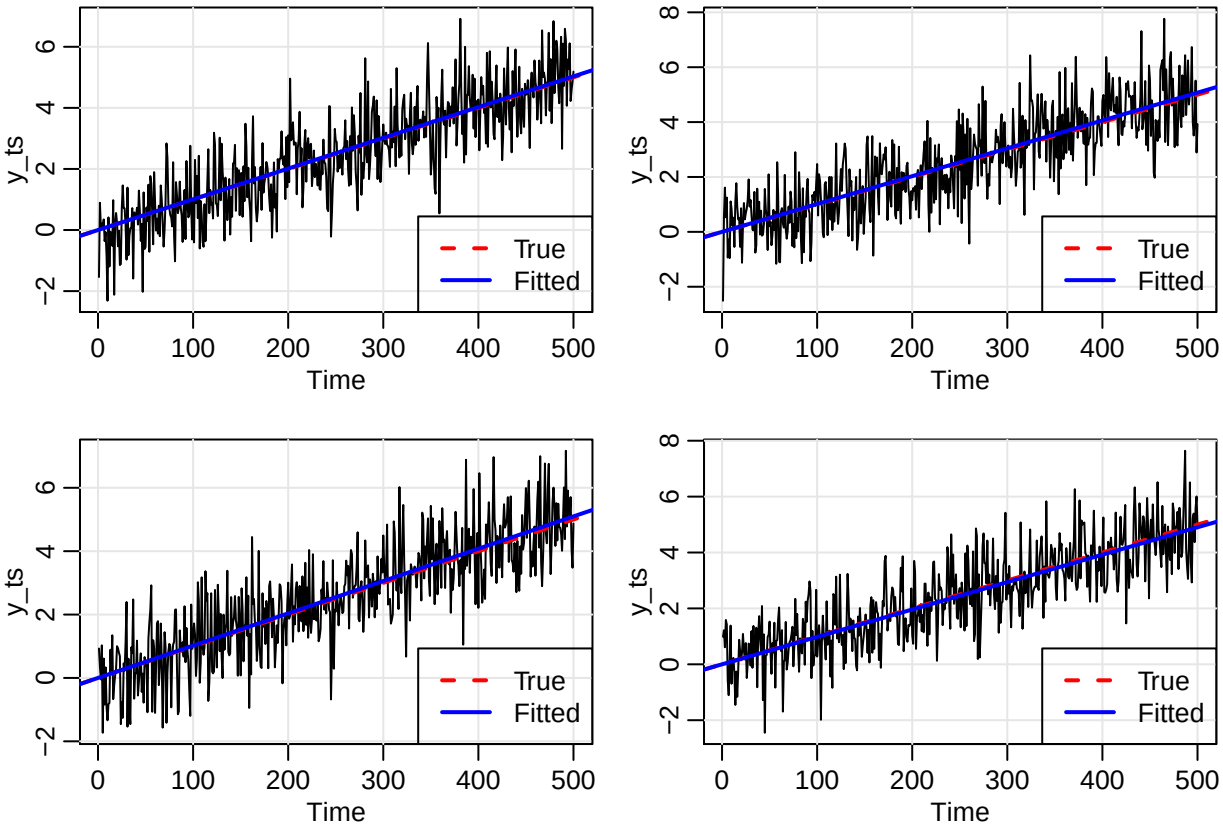



### Question 2.3.B

```
par(mfrow=c(2,2), mar=c(2.5,2.5,0,0)+.5, mgp=c(1.6,.6,0))

for (i in 1:4) {
  y <- numeric(500)
  t <- seq(1, 500, by=1)
  for (j in 1:500) {
   y[j] <- 0.01*t[j] + rnorm(1, 0, 1)
  }
  y_ts <- ts(y)
  regy <- lm(y_ts ~ 0 + time(y_ts))

  tsplot(y_ts)
  abline(a=0, b=0.01, col="red", lty=2, lwd=2)
  abline(regy, col="blue", lwd=2)
  legend("bottomright", c("True", "Fitted"), lty=c(2,1), lwd=c(2,2), col=c("red", "blue"))
}
```




### Question 2.3.C

From the images, we can see that the random walk with drift (first set of graphics) can stray far from the original mean function, because randomness is dictating the macro direction of the time series. The linear trends plus noise (second set of graphics) stick much closer to the mean function, because the noise only inserts minor variations, and does not dictate the macro trend.