Homework 2

Tony Silva

January 25, 2018

Number 1

Modules were read and watched

Number 2

The below R code generates a 100X100 matrix where each column is a new sample of 100 from the Gaussian distribution, N(0,1). In the code, I take the average of each column which provides the individual estimates, x_bar_i. Next, I find the variance of each of the columns which provides the individual estimates, sigma_hat_i.

In the code, these are desribed by the variables, x_bars and sigma_hats. Both are vectors of size 100.

```r
set.seed(5)
# Create a 100x100 matrix, each column is a new iid N(0,1)
x <- matrix(ncol = 100, nrow=100)
for (i in 1:100) {
  x[,i] <- rnorm(100,0,1)
}

#Obtain the all the sample means and variances.
#colMeans takes mean of each column and puts it in a vector
#apply gets all the column variances.
#This is the x_bar_i
x_bars <- colMeans(x)
sigma_hats <- apply(x, 2, var)

# Mean and Variance of the averages from each of the 100 experiments taking 100 samples from N(0,1)
mean(x_bars)
```

```
## [1] 0.001810254
```

```r
var(x_bars)
```

```
## [1] 0.009642371
```

```r
# Mean and Variance of the variances from each of the 100 experiments taking 100 samples from N(0,1)
mean(sigma_hats)
```

```
## [1] 1.02525
```

```r
var(sigma_hats)
```

```
## [1] 0.02048636
```

From the above calculations we can see the average of all of the x_bars is very close to 0. Additionally, the variance of the all of the x_bars is extremly small as well. The average for all of the sigma_hats is extremely close to 1. The variance for these values is extremely small as well. This is makes sense because we pulled our samples from the Gaussian Distribution.

Number 3

Let eps_t (Epsilon t) be the iid of N(0,1) for $1 <= t <= 500$

We define sigma_t as follows.

```
eps_t <- rnorm(500, 0,1)
```

Next, let's define x_t as x_t = eps_t + .5*eps_t-1 Assume, eps_t at t = 0 is 0.
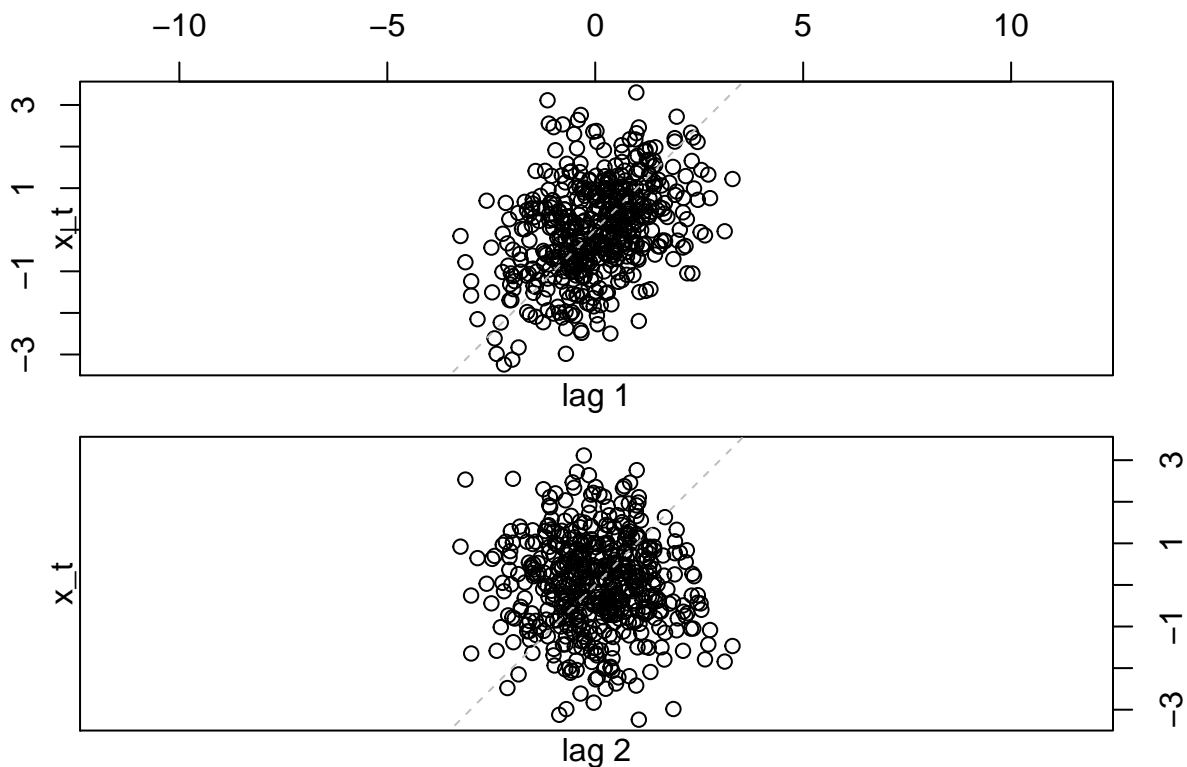
```
x_t <- vector(length=500)
for (i in 1:length(eps_t)) {
  if (i == 1) {
    x_t[i] <- eps_t[i]
  }else{
    x_t[i] <- eps_t[i] + .5*eps_t[i-1]
  }
}
```

We now have x_t. Now, we plot x_t vs. x_t-1, x_t-2, and t.

x_t vs. x_t-2

We can see the two lag plots below. The first is x_t vs. x_t-1. The second one is x_t vs. x_t-2

```
lag.plot(x_t, lags=2)
```


From the first lag plot, it appears there is correlation between the current value of x_t vs. the previous value of x_t (t-1).

x_t vs. t

Now, we plot x_t across time. From the plot below we can see that there is not necessarily a trend. However, what is interesting is when points appear across time they generally appear with one other point close by. This could be the correlation we were seeing in the plot of x_t vs. x_t-1.

```r
t <- 1:500
plot(x=t,y=x_t)
```
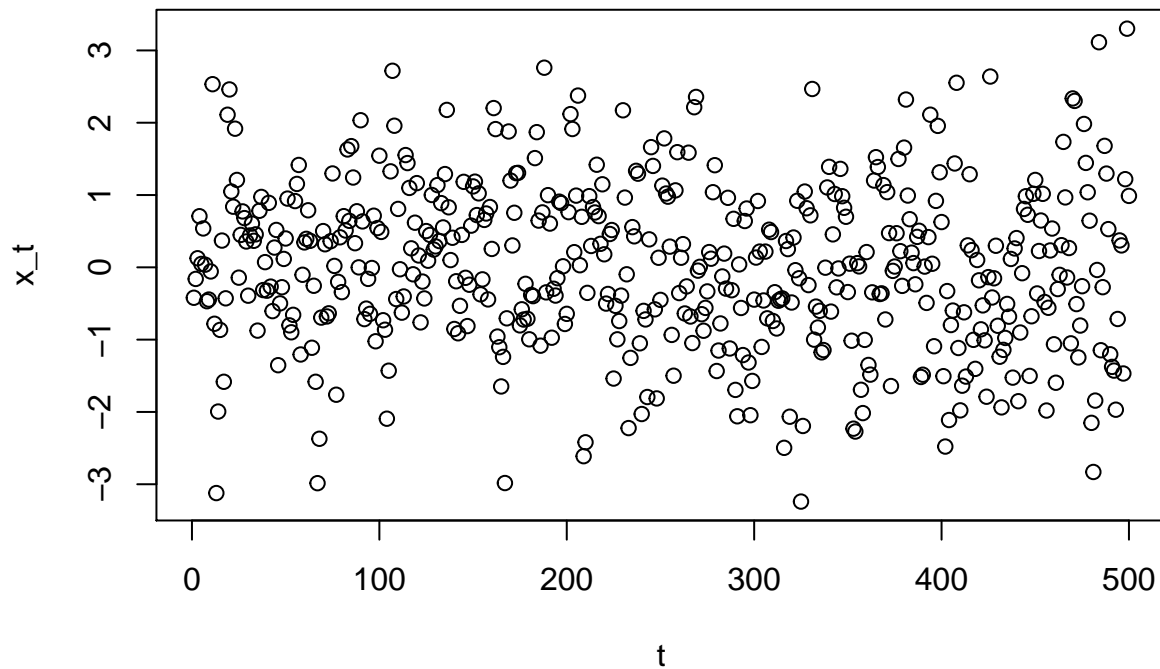

Mean, Variance, and Autocorrelation

Below is the average of x_t
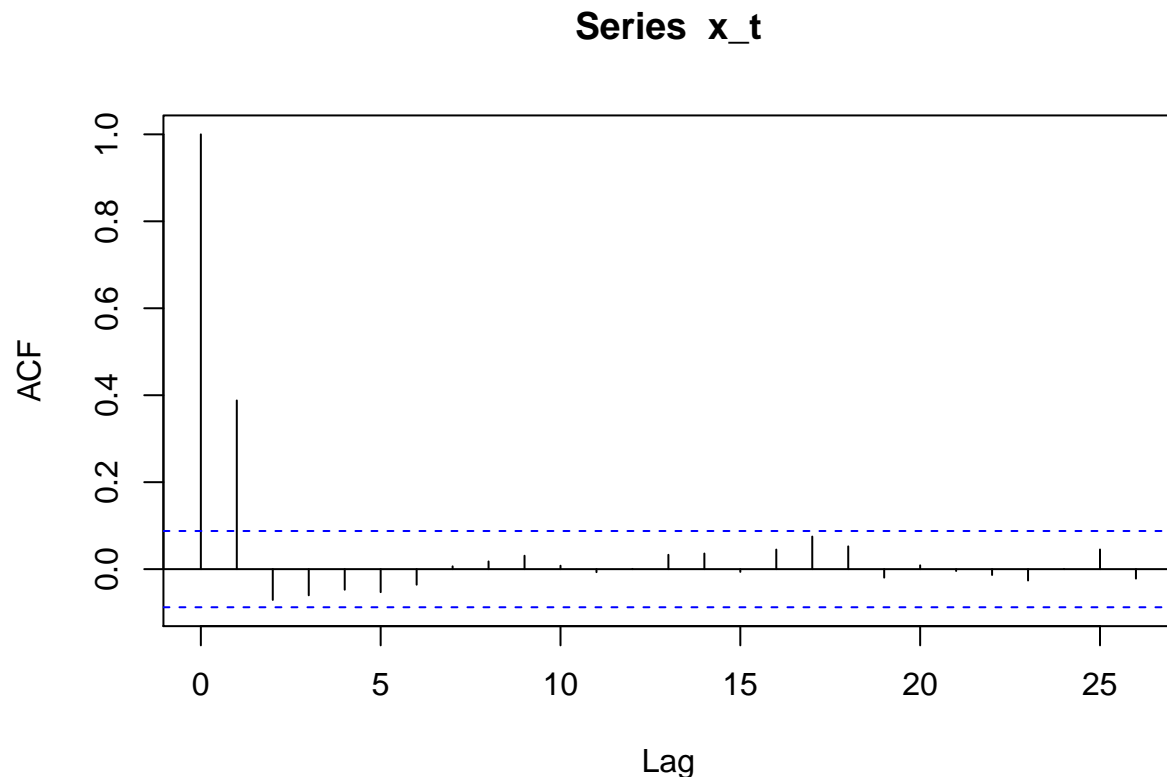
```r
mean(x_t)
```

```
## [1] 0.01978762
```

Below is the mean of x_t

```r
var(x_t)
```

```
## [1] 1.275798
```

Below is the autocorrelation of x_t and its plot

```r
acf(x_t, type = "correlation")
```

Series x_t

We can see from the Autocorrelation plot that x_t has a moderate correlation with its lag value of k=1. Which clarifies what we were seeing in the first plot and in the plot of x_t across t.

Number 4

I created a new vector for each of the normally distributed random variables with their differing variance. I then combined them into a R data frame and plotted the variables on the same plot using reshape2 and ggplot2 packages. The names of the variables are associated with their variance measure. On the plot I used a different color for each of the variables.

You can see the plot below.

```r
library(reshape2)
library(ggplot2)
x_.1 <- rnorm(100,0,.1)
x_.5 <- rnorm(100,0,.5)
x_1 <- rnorm(100,0,1)
x_2 <- rnorm(100,0,2)
x_5 <- rnorm(100,0,5)

data <- as.data.frame(cbind(x_.1,x_.5,x_1,x_2,x_5))
data.m <- melt(data)

## No id variables; using all as measure variables

g<-ggplot(data.m, aes(value, color = variable)) + geom_density()
g<- g+ggtitle("Plot of Multiple Normal Distributions with Different Variance")
```

Plot of Multiple Normal Distributions with Different Variance
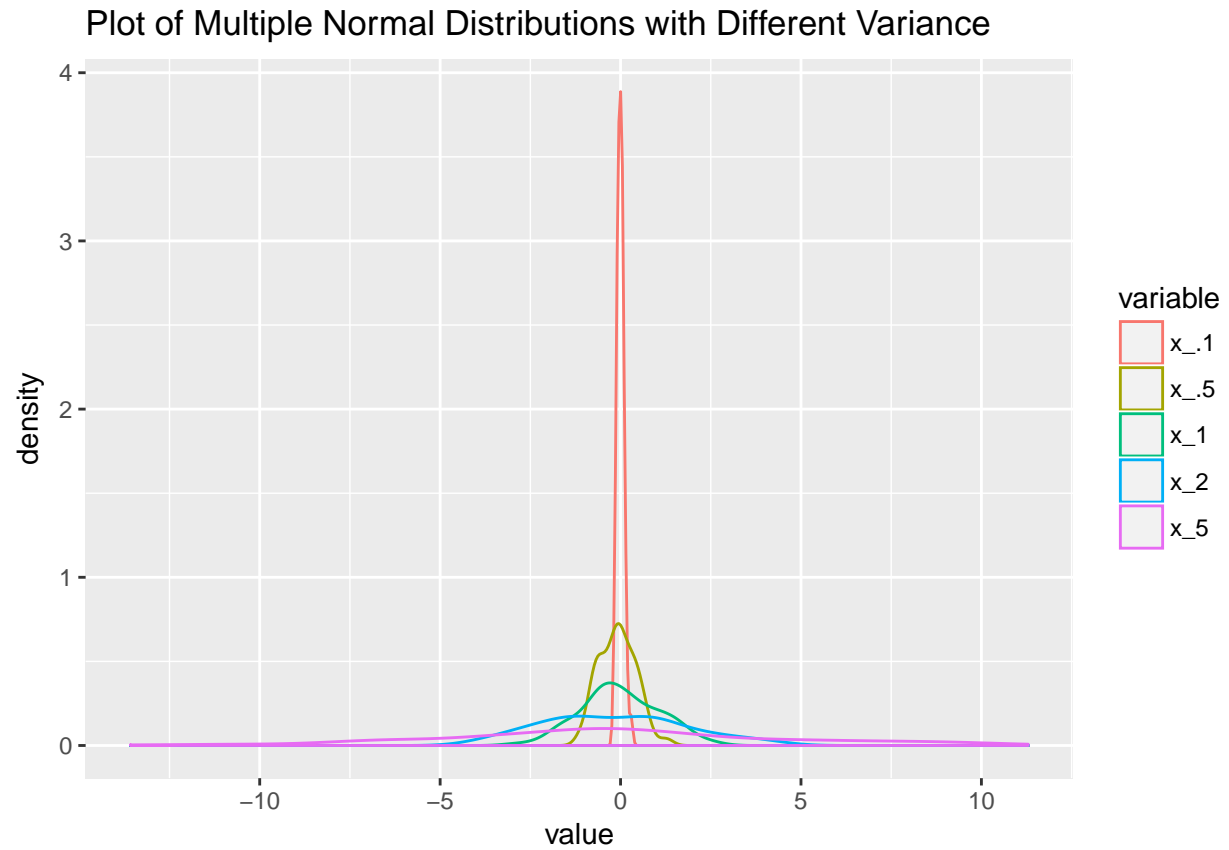
From the plot, you can see as the variance increases the curve grows wider and wider. As the variance gets smaller the variance gets smaller and smaller.