# Project_CLT.R

Salvador

Tue Jun 30 12:18:37 2015

```r
# Central Limit Theorem Project
# Our objective is to demonstrate the Central Limit Theorem
# by taking samples of increasing size out of a normal population
# We demonstrate that the sample mean is an unbiased estimator
# of the population mean by increasing sample size

mu <- 54
sigma <- 3

#sample of size 1
B <- 1000
s1 <- rnorm(n=B,m=mu,sd=sigma)
B <- 10000
s2 <- rnorm(n=B,m=mu,sd=sigma)
B <- 1000000
s3 <- rnorm(n=B,m=mu,sd=sigma)

attach(mtcars)
par(mfrow=c(3,1),mar=rep(2,4))
hist(s1,prob=TRUE,main="1000 Samples of Size 1",col="green",xlim=c(40,65))
curve(dnorm(x,mean=mean(s1),sd=sd(s1)),add=TRUE,col="red",lwd=2)
hist(s2,prob=TRUE,main="10000 Samples of Size 1",col="green",xlim=c(40,65))
curve(dnorm(x,mean=mean(s2),sd=sd(s2)),add=TRUE,col="red",lwd=2)
hist(s3,prob=TRUE,main="1000000 Samples of Size 1",col="green",xlim=c(40,65))
curve(dnorm(x,mean=mean(s3),sd=sd(s3)),add=TRUE,col="red",lwd=2)
```

1000 Samples of Size 1

10000 Samples of Size 1

1000000 Samples of Size 1
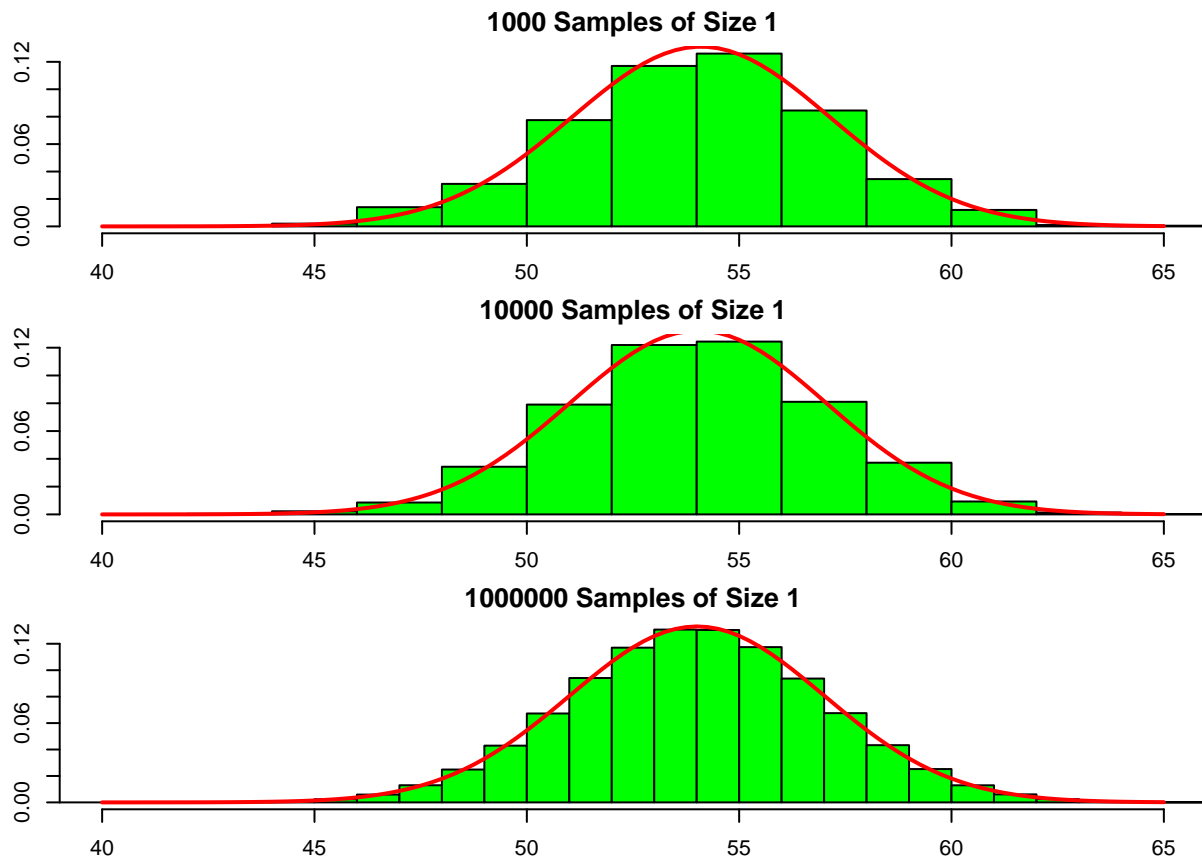
```
c(mean(s1),sd(s1))
```

```
## [1] 54.09941   3.04419
```

```
c(mean(s2),sd(s2))
```

```
## [1] 54.032689   3.014948
```

```
c(mean(s3),sd(s3))
```

```
## [1] 54.006259   2.997538
```

```
#Samples of size n=2
n <- 2

B <- 1000
s1_2 <- numeric(B)
for ( i in 1:B ) {s1_2[i]=mean(rnorm(n=n,m=mu,sd=sigma))}

B <- 10000
s2_2 <- numeric(B)
for ( i in 1:B ) {s2_2[i]=mean(rnorm(n=n,m=mu,sd=sigma))}
```
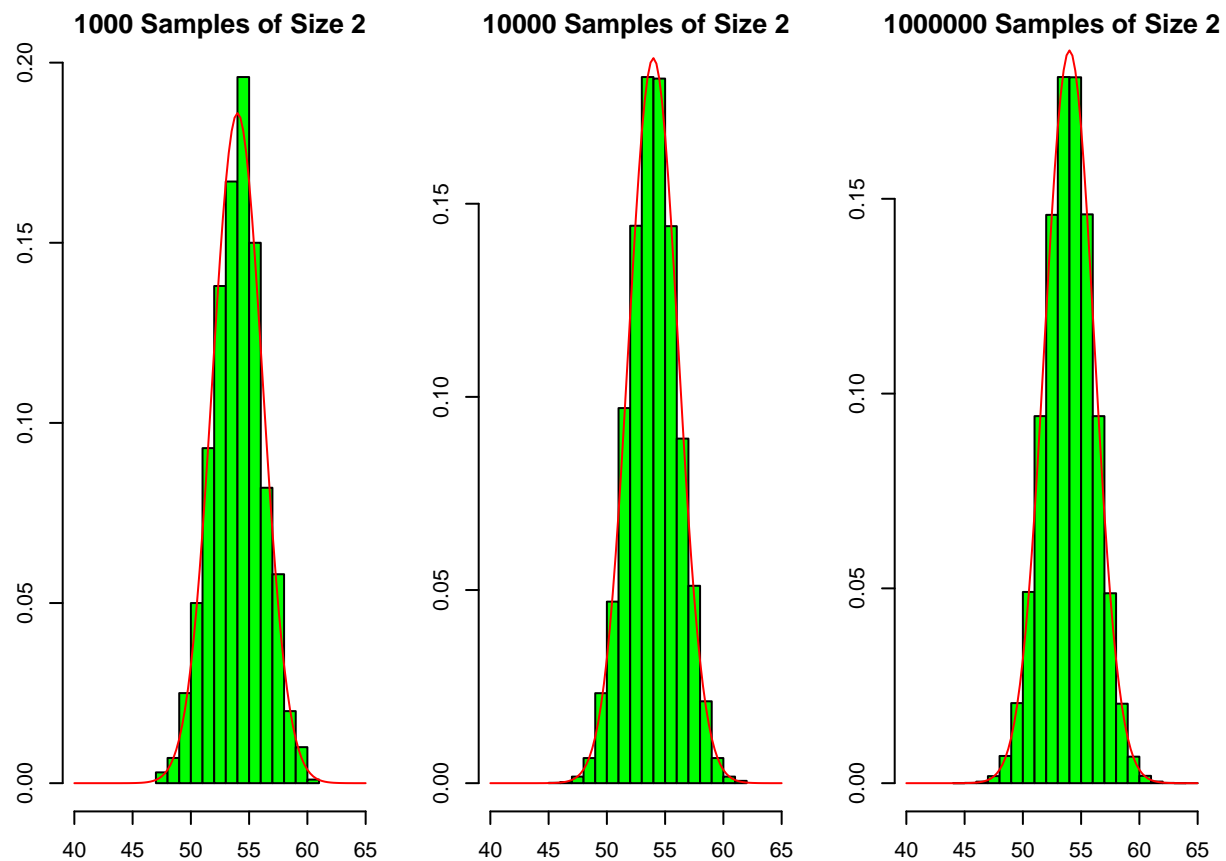
```
B <- 1000000
s3_2 <- numeric(B)
for ( i in 1:B ) {s3_2[i]=mean(rnorm(n=n,m=mu,sd=sigma))}

par(mfrow=c(1,3),mar=rep(2,4))
hist(s1_2,prob=TRUE,main="1000 Samples of Size 2",col="green",xlim=c(40,65))
curve(dnorm(x,mean=mean(s1_2),sd=sd(s1_2)),add=TRUE,col="red")
hist(s2_2,prob=TRUE,main="10000 Samples of Size 2",col="green",xlim=c(40,65))
curve(dnorm(x,mean=mean(s2_2),sd=sd(s2_2)),add=TRUE,col="red")
hist(s3_2,prob=TRUE,main="1000000 Samples of Size 2",col="green",xlim=c(40,65))
curve(dnorm(x,mean=mean(s3_2),sd=sd(s3_2)),add=TRUE,col="red")
```



```
c(mean(s1_2),sd(s1_2))
```

```
## [1] 54.000976   2.145082
```

```
c(mean(s2_2),sd(s2_2))
```

```
## [1] 53.987127   2.125353
```

```
c(mean(s3_2),sd(s3_2))
```

```
## [1] 53.997625   2.120987
```
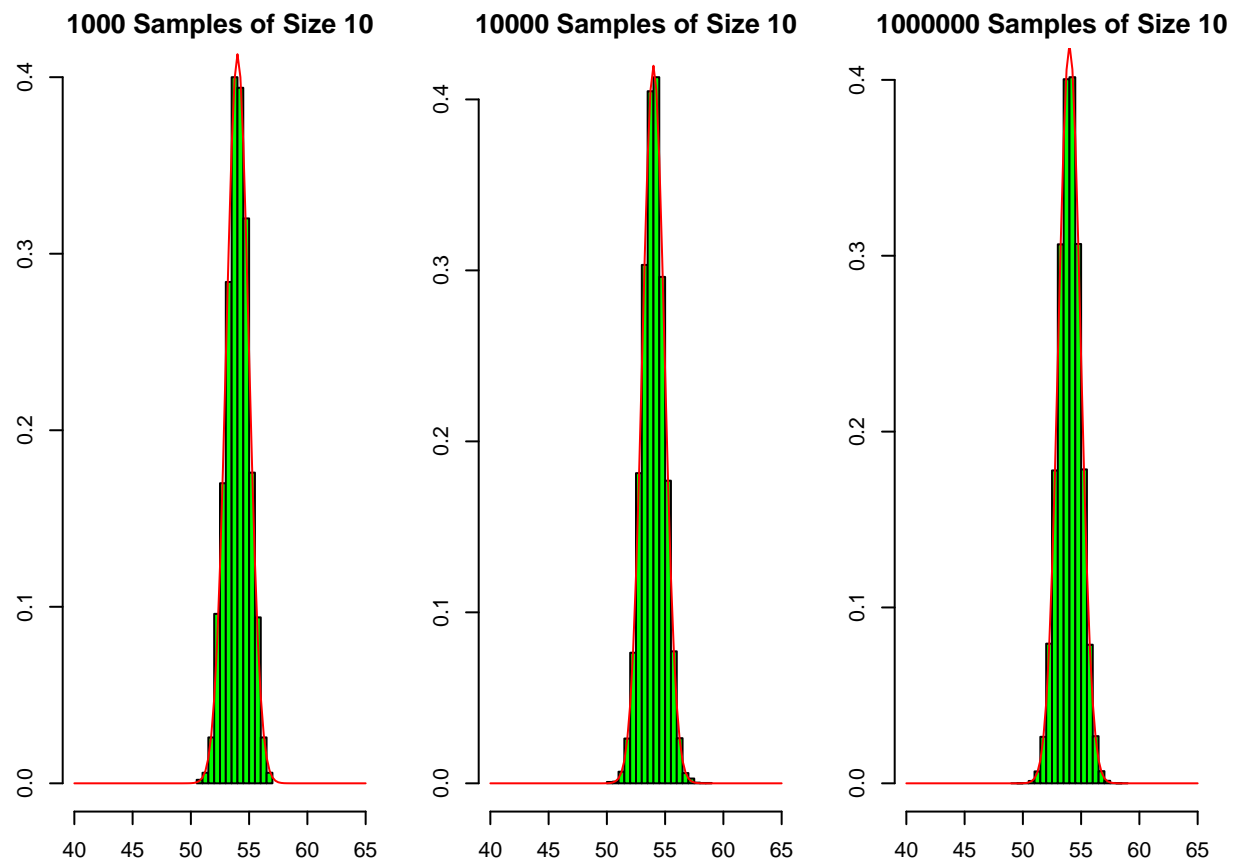
```
#Samples of size n=10
n <- 10

B <- 1000
s1_10 <- numeric(B)
for ( i in 1:B ) {s1_10[i]=mean(rnorm(n=n,m=mu,sd=sigma))}

B <- 10000
s2_10 <- numeric(B)
for ( i in 1:B ) {s2_10[i]=mean(rnorm(n=n,m=mu,sd=sigma))}

B <- 1000000
s3_10 <- numeric(B)
for ( i in 1:B ) {s3_10[i]=mean(rnorm(n=n,m=mu,sd=sigma))}

par(mfrow=c(1,3),mar=rep(2,4))
hist(s1_10,prob=TRUE,main="1000 Samples of Size 10",col="green",xlim=c(40,65))
curve(dnorm(x,mean=mean(s1_10),sd=sd(s1_10)),add=TRUE,col="red")
hist(s2_10,prob=TRUE,main="10000 Samples of Size 10",col="green",xlim=c(40,65))
curve(dnorm(x,mean=mean(s2_10),sd=sd(s2_10)),add=TRUE,col="red")
hist(s3_10,prob=TRUE,main="1000000 Samples of Size 10",col="green",xlim=c(40,65))
curve(dnorm(x,mean=mean(s3_10),sd=sd(s3_10)),add=TRUE,col="red")
```



```
c(mean(s1_10),sd(s1_10))
```

```
## [1] 54.0054105  0.9658525
```

```
c(mean(s2_10),sd(s2_10))
```

```
## [1] 53.9980166  0.9505617
```

```
c(mean(s3_10),sd(s3_10))
```

```
## [1] 54.0008018  0.9494492
```

```
#Samples of size n=30
n <- 30

B <- 1000
s1_30 <- numeric(B)
for ( i in 1:B ) {s1_30[i]=mean(rnorm(n=n,m=mu,sd=sigma))}

B <- 10000
s2_30 <- numeric(B)
for ( i in 1:B ) {s2_30[i]=mean(rnorm(n=n,m=mu,sd=sigma))}

B <- 1000000
s3_30 <- numeric(B)
for ( i in 1:B ) {s3_30[i]=mean(rnorm(n=n,m=mu,sd=sigma))}

par(mfrow=c(1,3),mar=rep(2,4))
hist(s1_30,prob=TRUE,main="1000 Samples of Size 30",col="green",xlim=c(40,65))
curve(dnorm(x,mean=mean(s1_30),sd=sd(s1_30)),add=TRUE,col="red")
hist(s2_30,prob=TRUE,main="10000 Samples of Size 30",col="green",xlim=c(40,65))
curve(dnorm(x,mean=mean(s2_30),sd=sd(s2_30)),add=TRUE,col="red")
hist(s3_30,prob=TRUE,main="1000000 Samples of Size 30",col="green",xlim=c(40,65))
curve(dnorm(x,mean=mean(s3_30),sd=sd(s3_30)),add=TRUE,col="red")
```

1000 Samples of Size 30   **10000 Samples of Size 30**   **1000000 Samples of Size 30**
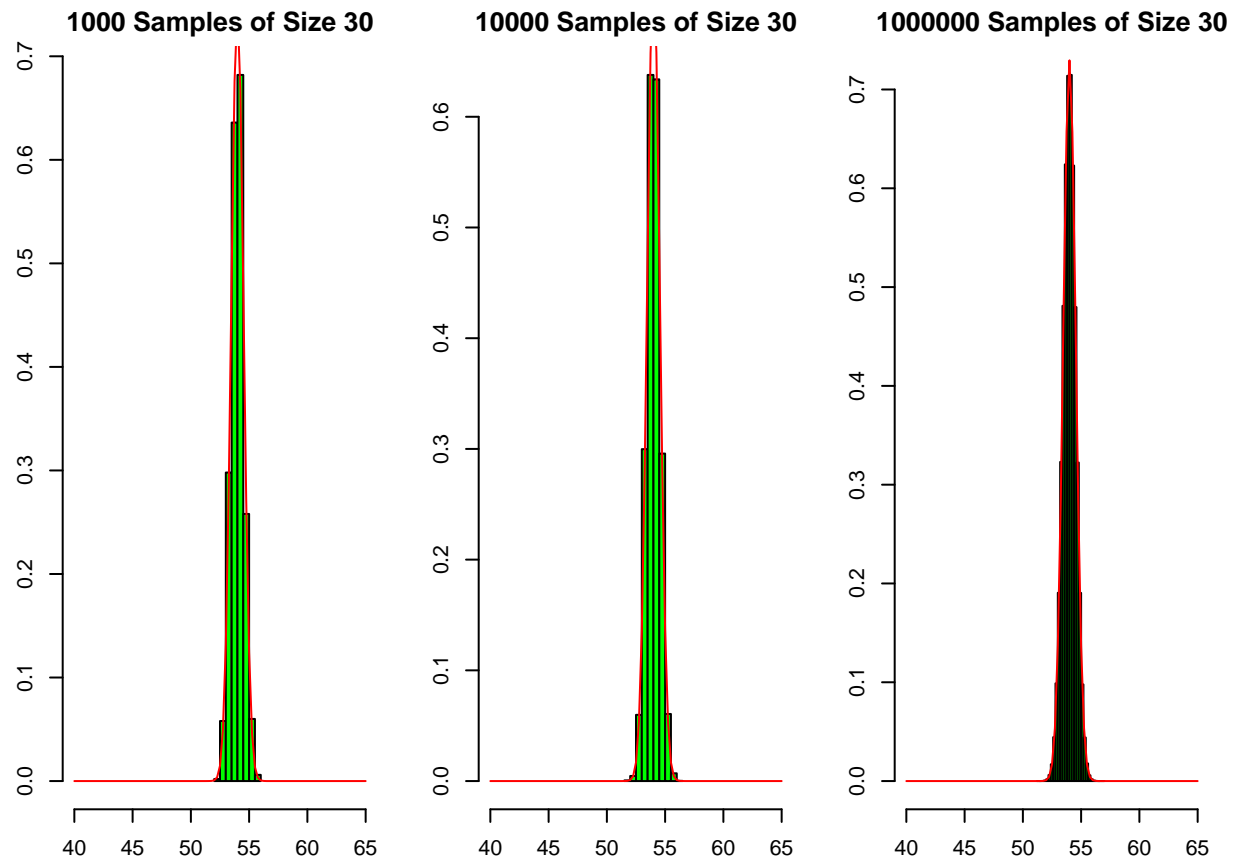
```r
c(mean(s1_30),sd(s1_30))
```

```
## [1] 53.9918116  0.5323285
```

```r
c(mean(s2_30),sd(s2_30))
```

```
## [1] 54.0003194  0.5460199
```

```r
c(mean(s3_30),sd(s3_30))
```

```
## [1] 53.9995346  0.5468374
```

```r
detach(mtcars)
```