# Statistical Inference Course Project Author: Anonymous

The purpose of this project is to illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponentials. Namely, 1. To show the sample mean and compare it to the theoretical mean of the distribution. 2. To show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution. 3. To show that the distribution is approximately normal.

First, I will perform the simulation

```
set.seed(65536)
mns = NULL
for (i in 1 : 10000) mns = c(mns, mean(rexp(40,0.2)))
```

Remember, when lambda=0.2, the theoretical mean and stdev of the exponential distribution are equal to 1/0.2=5

According to CLT, the average of random samples from a distribution will follow the normal distribution with mean=mean of the distribution and var=var of the distribution/number of samples being averaged. Hence, in theory, **mean(mns)=5**, **var(mns)=5^2/40=0.625**. Let's calculate the actual value from the simulation.

```
mean(mns)
```

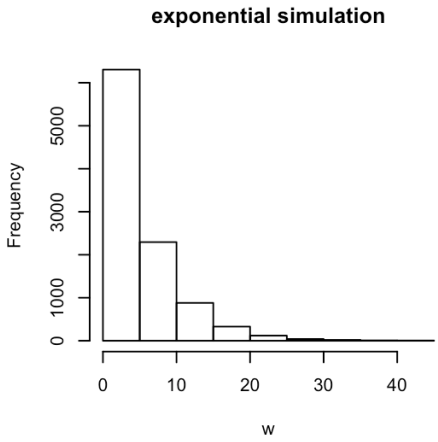
```
## [1] 5.001474
```

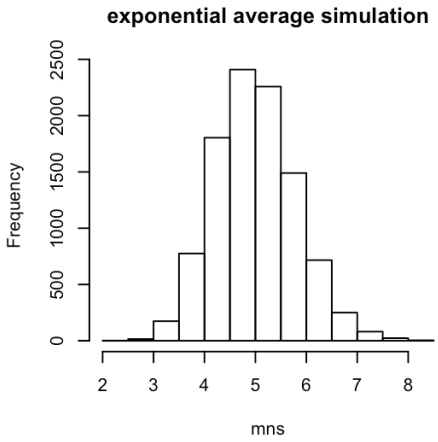
```
var(mns)
```

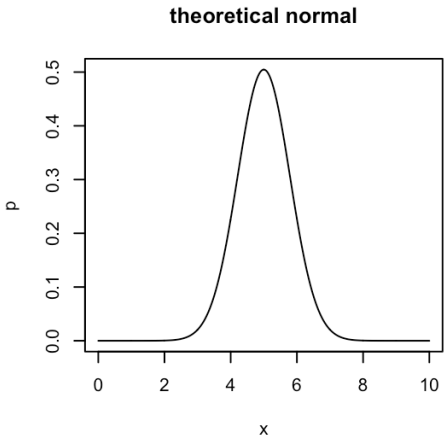
```
## [1] 0.6267727
```

Is the distribution similar to normal? Let's compare the distribution of simulation of samples, simulation of the 40 average of the samples, and a theoretical normal distribution

```
par(mfrow=c(1,3))
w=rexp(10000,0.2)
hist(w, main="exponential simulation")
hist(mns,main="exponential average simulation")
x=seq(0,10,length=1000)
y=dnorm(x,mean=5,sd=sqrt(25/40))
plot(x,y,type="l",main="theoretical normal",ylab="p")
```




It should tells well that the average is approximately normally distributed, compared with single randoms.