Coursera Statistical Inference Simulation Project Part 1 By: Hisam Sabouni

In this project I will be simulating the distribution of averages of 40 random exponential distributions with lambda = 0.2 for all simulations. I will show the following:

1. The sample mean and compare it to the theoretical mean of the distribution (1/lambda or in our case 1/0.2 = 5)
2. How variable the sample is (via variance) and compare it to the theoretical variance of the distribution
3. That the distribution is approximatley normal

Simulations: Here we have 10,000 simulations of the mean of 40 random exponentials with lambda = 0.2:

```
means.exp = NULL
var.exp = NULL
set.seed(10)
for (i in 1:10000){
  means.exp = c(means.exp, mean(rexp(40,0.2)))
  var.exp=c(var.exp, var(rexp(40,0.2)))
}
```

Sample Mean versus Theoretical Mean: A view of the data generated and some summary statistics of the distributions as a result of the simulations:

```
head(means.exp)
```

```
## [1] 4.832621 4.556120 5.513666 4.994439 3.674707 3.803456
```

```
length(means.exp)
```

```
## [1] 10000
```

```
summary(means.exp)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.610   4.462   4.970   5.007   5.515   8.385
```

The mean is approximatley 5 which is what the theoretical mean of the distribution should be.

Variance versus Thoretical Variance: Now a look at the mean variance of the 10,000 samples of 40:

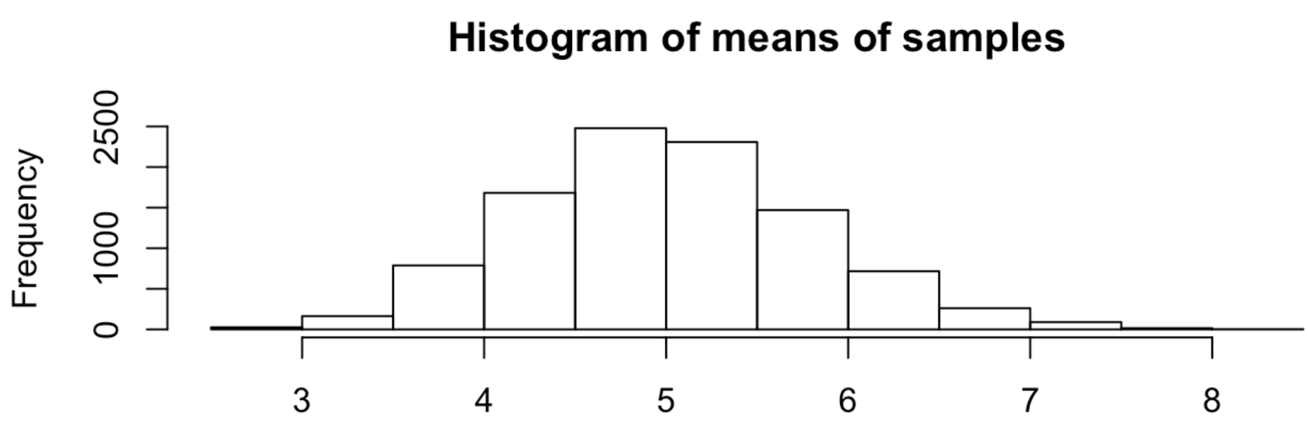
```
mean(var.exp)
```

```
## [1] 24.95536
```

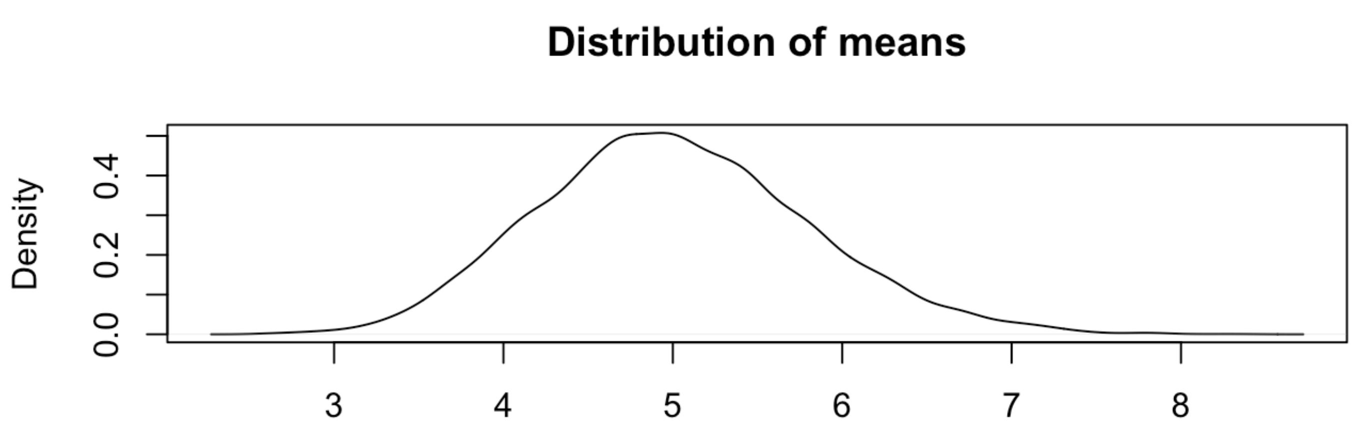
The mean of the variance is approximatley 25, which coincides with the theoretical variance of the distribution.

Plots: A view of the distribution of samples:

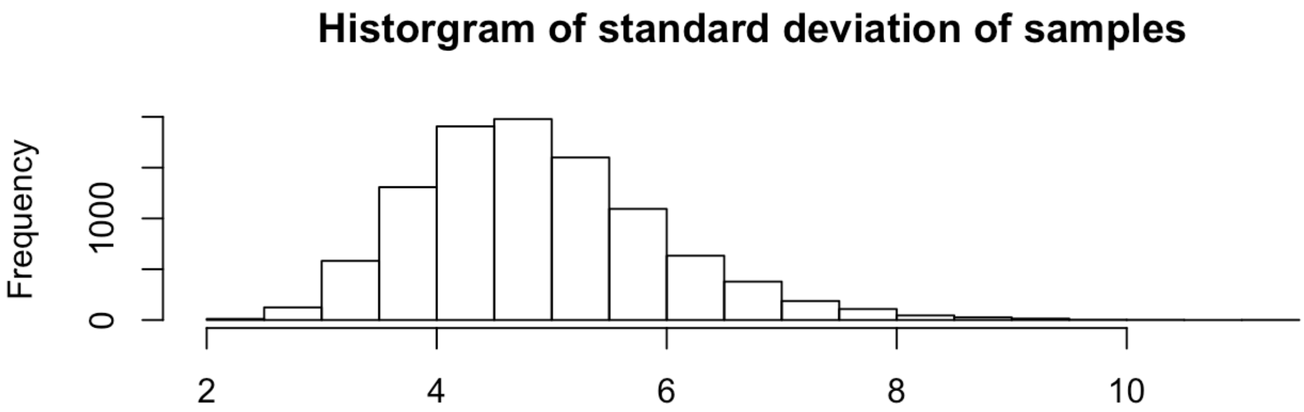
```
hist(means.exp,main="Histogram of means of samples",xlab="")
```




```
plot(density(means.exp),main="Distribution of means",xlab="")
```




```
hist(sqrt(var.exp),main = "Historgram of standard deviation of samples",xlab="")
```



As you can see the density of the means of the random samples is approximatley normal, thanks to the Central Limit Theorem. We can look at this in more detail by running the following code:

```
density(means.exp)
```

```
##
## Call:
##  density.default(x = means.exp)
##
## Data: means.exp (10000 obs.);    Bandwidth 'bw' = 0.1122
##
##        x               y
##  Min.   :2.274   Min.   :0.0000042
##  1st Qu.:3.886   1st Qu.:0.0041336
##  Median :5.498   Median :0.0637564
##  Mean   :5.498   Mean   :0.1549378
##  3rd Qu.:7.110   3rd Qu.:0.3000792
##  Max.   :8.722   Max.   :0.5074801
```

Notice how the mean and median are approximatley the same, this tells us that we have a symmetric distribution (like the normal distribution)