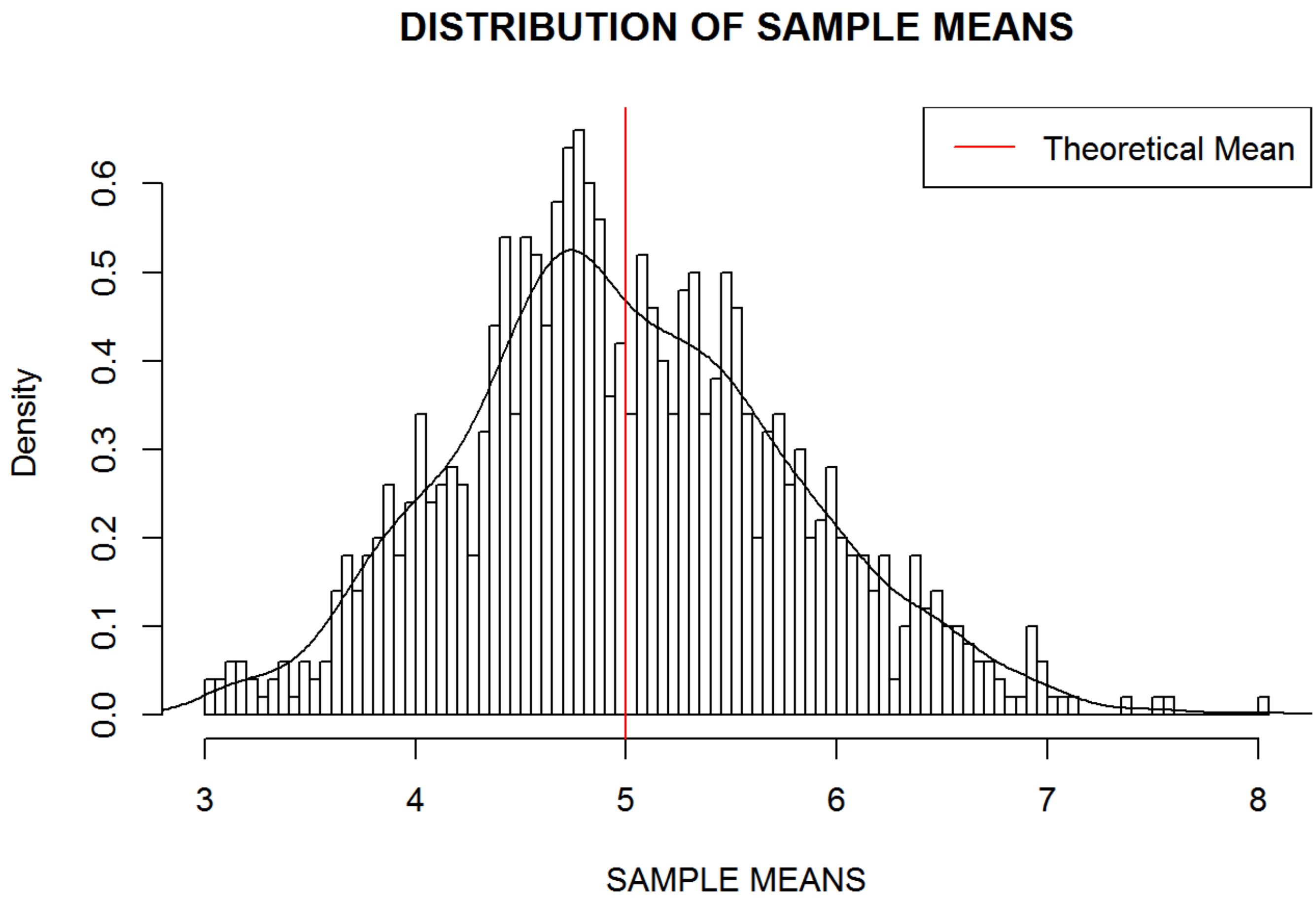
Coursera_DS_Inference_Project1

Atul

January 31, 2016

In this project we investigate the exponential distribution in R and compare it with the Central Limit Theorem. The exponential distribution can be simulated in R with rexp(n, lambda) where lambda is the rate parameter. The mean of exponential distribution is 1/lambda and the standard deviation is also 1/lambda. Set lambda = 0.2 for all of the simulations. We investigate the distribution of averages of 40 exponentials. We do a thousand simulations here.

```
set.seed(100)
mns = NULL
n <- 40
lambda <- 0.2
for (i in 1:1000) {mns = c(mns, mean(rexp(n,rate = lambda)))}
hist(mns, prob = T, xlab = "SAMPLE MEANS", main = "DISTRIBUTION OF SAMPLE MEANS", breaks=
100)
lines(density(mns))
abline(v=1/lambda, col="red")
legend('topright', "Theoretical Mean", lty=1, col="red")
```




```
print(paste("Theoretical Mean =", 1/lambda))
```

```
## [1] "Theoretical Mean = 5"
```

```
print(paste("Mean of sample means = ",mean(mns)))
```

```
## [1] "Mean of sample means =  4.9997019268744"
```

In above figure, we have produce sample means of size 40, we see that mean of samples is very close to theretical mean by central limit theorem.

```
print(paste("THEORETICAL STD DEVIATION =", (1/lambda)/sqrt(n)))
```

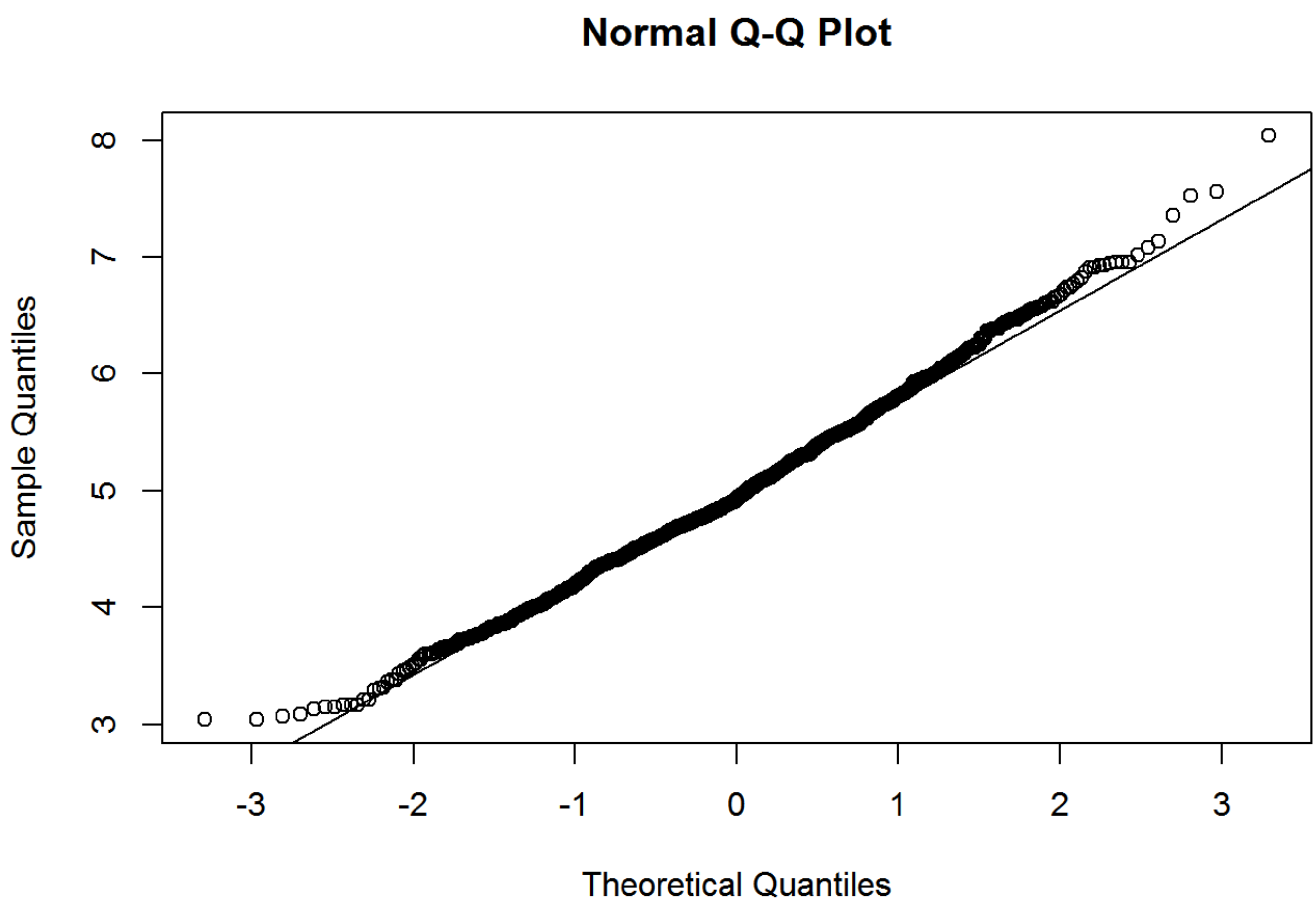
```
## [1] "THEORETICAL STD DEVIATION = 0.790569415042095"
```

```
print(paste("STD DEVIATION OF SIMULATED MEANS =", sd(mns)))
```

```
## [1] "STD DEVIATION OF SIMULATED MEANS = 0.802025077464672"
```

Above calculations show that sample means' variance like mean is also close to variance by central limit theorem.

```
qqnorm(mns)
qqline(mns)
```



As illustrate in above Q-Q plot, quantiles of normal and sample distribution closely match. This proves validity of normal distribution hypothesis.