Central Limit Theorem

Dongli Liu

2015 9 23

This document is to investigate the exponential distribution in R and compare it with the Central Limit Theorem. As requested by the instructor, pre-conditions are:
1. Set *lambda = 0.2* for all of the simulations;
2. Investigate the distribution of averages of 40 exponentials;
3. do a thousand simulations.

Illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponentials. We need to:
1. Show the sample mean and compare it to the theoretical mean of the distribution;
2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution;
3. Show that the distribution is approximately normal.

Here we do this step by step. Firstly let's set *lambda = 0.2*

```
lambda <- 0.2
```

Then theoretically, for a distribution of averages of 40 exponentials, mean is:

```
1/lambda;
```

```
## [1] 5
```

standard deviation is:

```
(1/lambda)/(40)^0.5;
```

```
## [1] 0.7905694
```

Now is the simulation parts:

```
set.seed(923) # So that we could reproduce this result
mns = NULL
for (i in 1: 1000) mns = c(mns, mean(rexp(40, rate=lambda)))
```

It is easy to see that sample mean is nearly same as theoretical mean:

```
mean(mns)
```

```
## [1] 4.987689
```

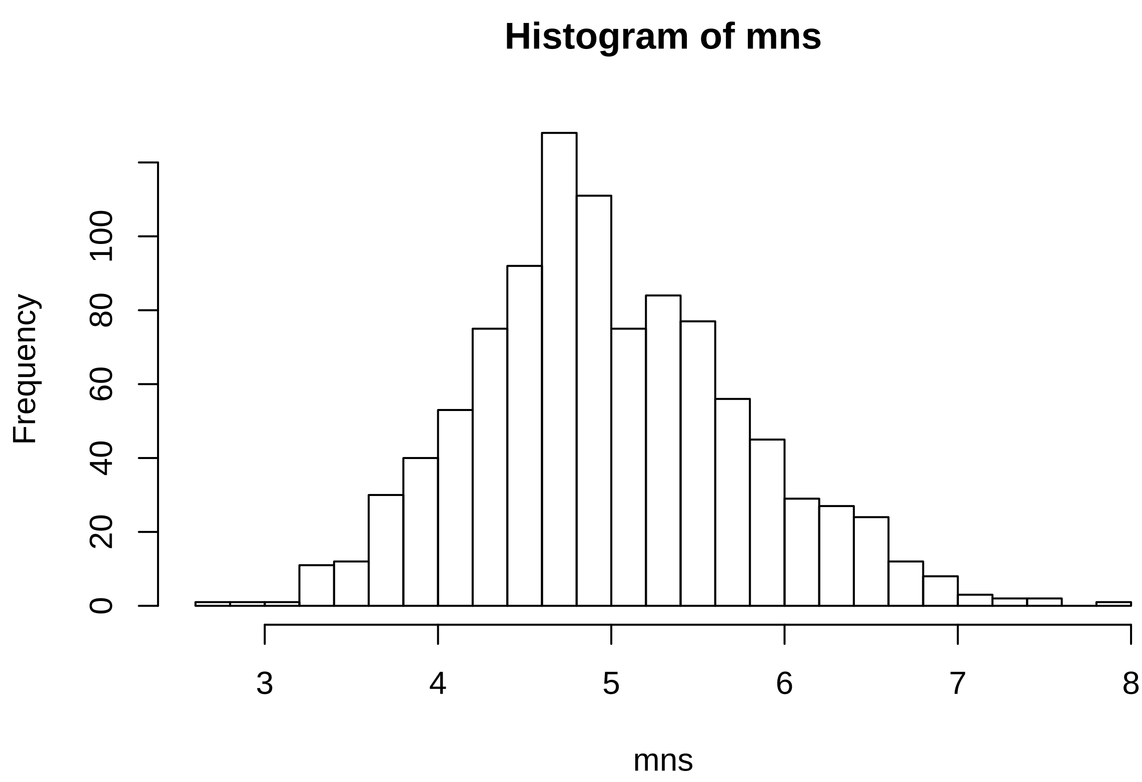
And for sample standard deviation:

```r
sd(mns)
```

```
## [1] 0.7890795
```

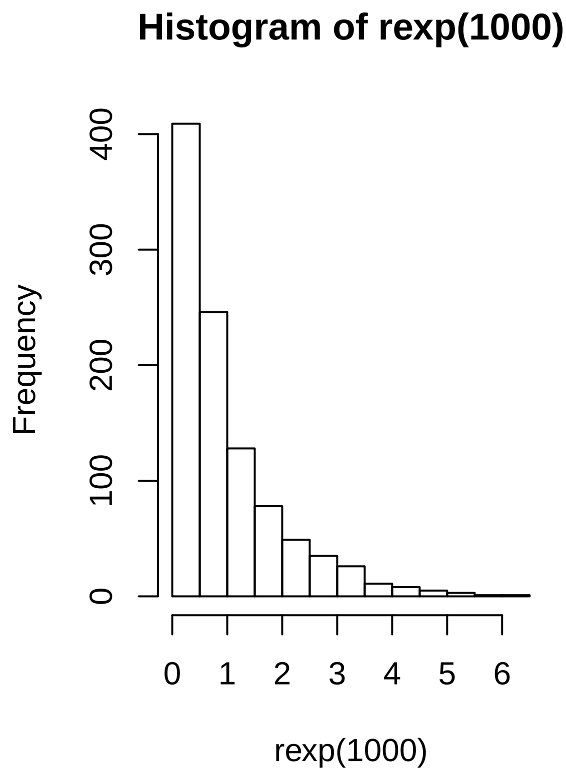
Next, let's have a look at the histogram of this distribution:

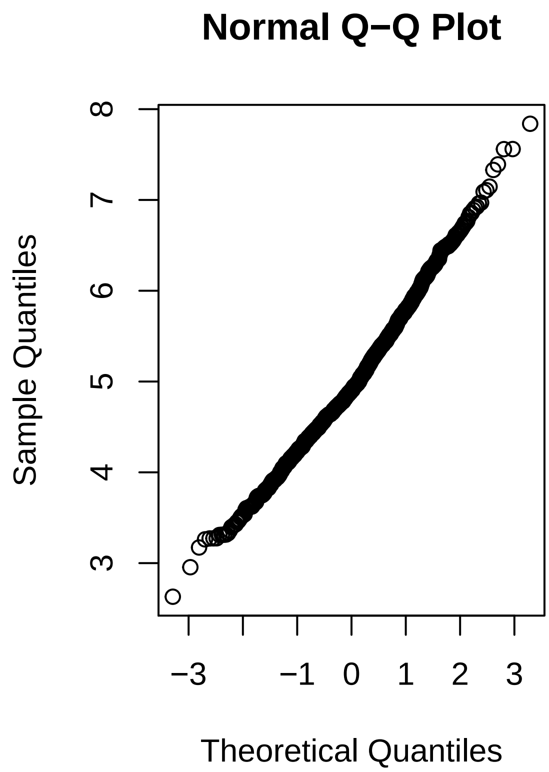
```r
hist(mns, breaks=20)
```



We could see that the histogram is quite different with distribution of 1000 random exponentials; instead, it is much more like a Gaussian distribution via normal Q-Q plot:

```r
par(mfrow = c(1,2))
hist(rexp(1000))
qqnorm(mns)
```


Histogram of rexp(1000)
Frequency
0
100
200
300
400
0 1 2 3 4 5 6
rexp(1000)


Normal Q−Q Plot
Sample Quantiles
3
4
5
6
7
8
−3 −1 0 1 2 3
Theoretical Quantiles