Simulation of the Central Limit Theorem

Sani Stephen

April 4, 2016

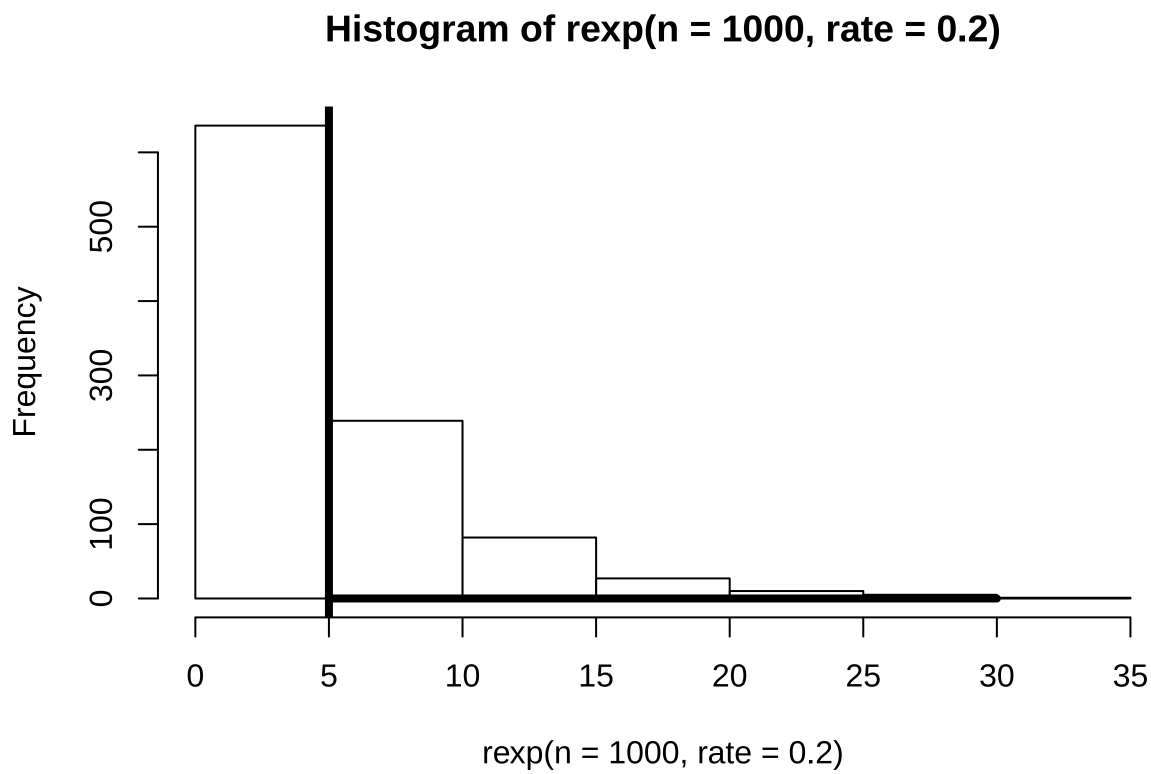
Overview

This project report is on the simulation of the Central Limit Theorem with the exponential distribution. A thausand averages of 40 exponentials with lambda (rate) = 0.2 was simulated. Theory: The mean and standard deviation of exponential distribution are both 1/lambda (rate).

Simulations

The exponential distribution for a thousand random variables with a lambda of 0.2 is shown below

```
hist(rexp(n = 1000, rate = 0.2))
abline(v=(1/0.2), lwd = 4)
segments(1/0.2,0,(1/0.2)+(1/0.2)^2,0,lwd = 4)
```



The distribution is rigth skewed. The vertical line through 1/0.2 or 5 is the theoretical mean ie Theoretical mean is 5 The Theoretical standard deviation is 5 (or 1/0.2) and The theoretical variance is 25 (or 1/0.2^2)

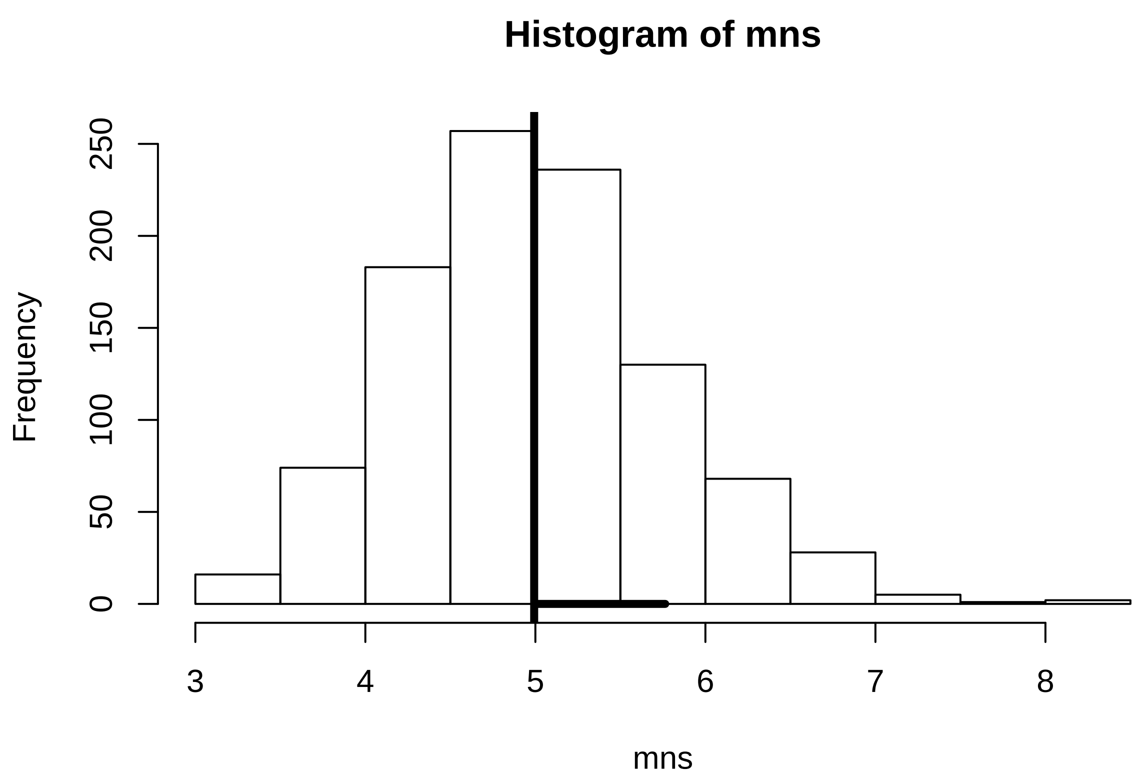
The horizontal line between 5 and 5 + (1/0.2) or 10 is the theoretical standard deviation around the theoretical mean

To simulate the Central Limit Theorem, the following assumptions are made

1. The random variables are independent and identically distributed
2. The number of samples is large enough: n > 30

The distribution of 1000 simulation of the mean of 40 random samples of the random variables is shown below

```
mns = NULL
for (i in 1 : 1000) mns = c(mns, mean(rexp(40, 0.2)))
hist(mns)
abline(v=mean(mns), lwd=4)
segments(mean(mns),0,mean(mns)+sd(mns),0,lwd = 4)
```



The distribution looks more normal The vertical line through

```
mean(mns)
```

```
## [1] 4.992866
```

is the mean of the sample mean The standard deviation of the sample means is

```
sd(mns)
```

```
## [1] 0.7709639
```

The horizontal line between mean and (mean + standard deviation) is the standard deviation of the sample means around the mean of the sample means

The variance of the sample means is

```
var(mns)
```

```
## [1] 0.5943853
```

Sample Mean versus Theoretical Mean

The sample mean and the theoretical mean are approximately equal which confirms the Central Limit Theorem.

Sample Variance versus Theoretical Variance

The standard deviation as well as the variance of the sample means are much less than those of the original distribution and thus less variable. Thus the mean of the sample mean is a very good approximation of the distribution mean.

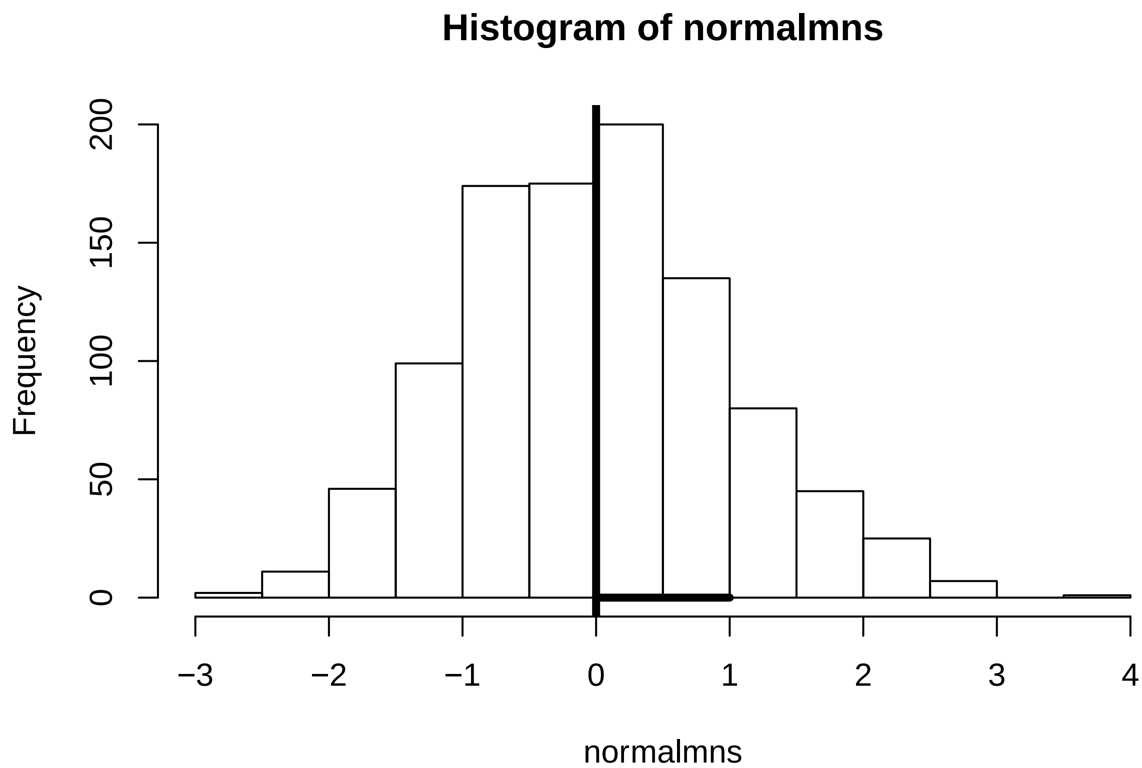
Distribution:

The exponential distribution for a thousand random variables with a lambda of 0.2 is shown is rigth skewed.

The distribution of 1000 simulation of the mean of 40 random samples of the random variables is looks more normal

The plot of the normalised sample meanes should have a mean of 0 and a standard deviation of 1

```
mns = NULL
for (i in 1 : 1000) mns = c(mns, mean(rexp(40, 0.2)))
normalmns <-(mns-mean(mns))/sd(mns)
hist(normalmns)
abline(v=mean(normalmns), lwd=4)
segments(mean(normalmns),0,mean(normalmns)+sd(normalmns),0,lwd = 4)
```



mean of the normalised distribution is

```r
mean(normalmns)
```

```
## [1] 3.684931e-16
```

standard deviation of the normalised distribution is

```r
sd(normalmns)
```

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