Central Limit Theorem

Michael McDermott

December 13, 2014

Central Limit Theorem in Action

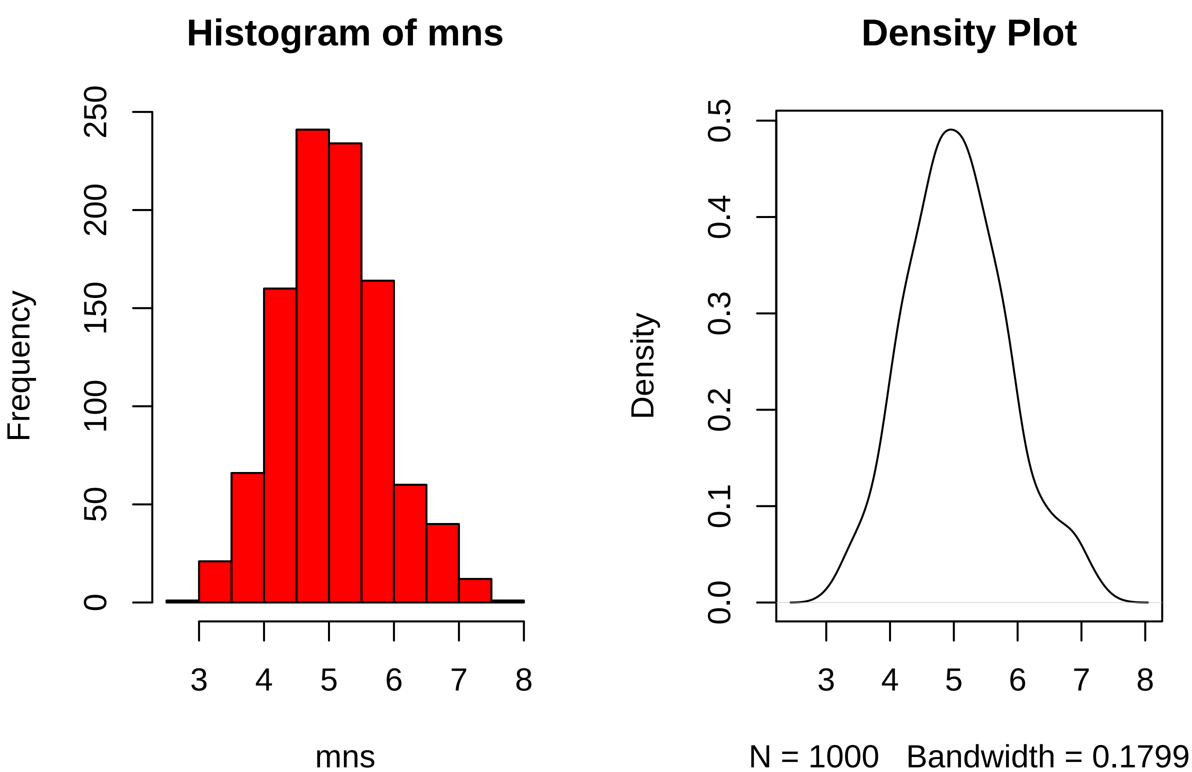
In this document I investigate the Central Limit Theorem for the exponential distribution.

Start by taking an exponential distribution with $\lambda = 0.2$ (and so average $1/\lambda$) and sampling it 40 times, then taking the mean, and repeating this process 1000 times:

```
library(ggplot2)
set.seed(1337)
mns = NULL
l=0.2
for (i in 1 : 1000) mns = c(mns, mean(rexp(40,l)))
```

The distribution of the averages is:

```
par(mfrow=c(1,2))
hist(mns,col='red')
plot(density(mns),main='Density Plot')
```



Now let's compute the average. The expected average of the average of distributions is the average of the distribution itself, in this case $1/\lambda = 5$.

```
expected_average = 1/l
expected_average
```

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

```
actual_average = mean(mns)
actual_average
```

```
## [1] 5.055995
```

Similarly the expected variance of the averages of a distribution is the variance of the original distribution divided by the sample size, in this case 40.

```
expected_var = (1/l)^2/(40)
expected_var
```

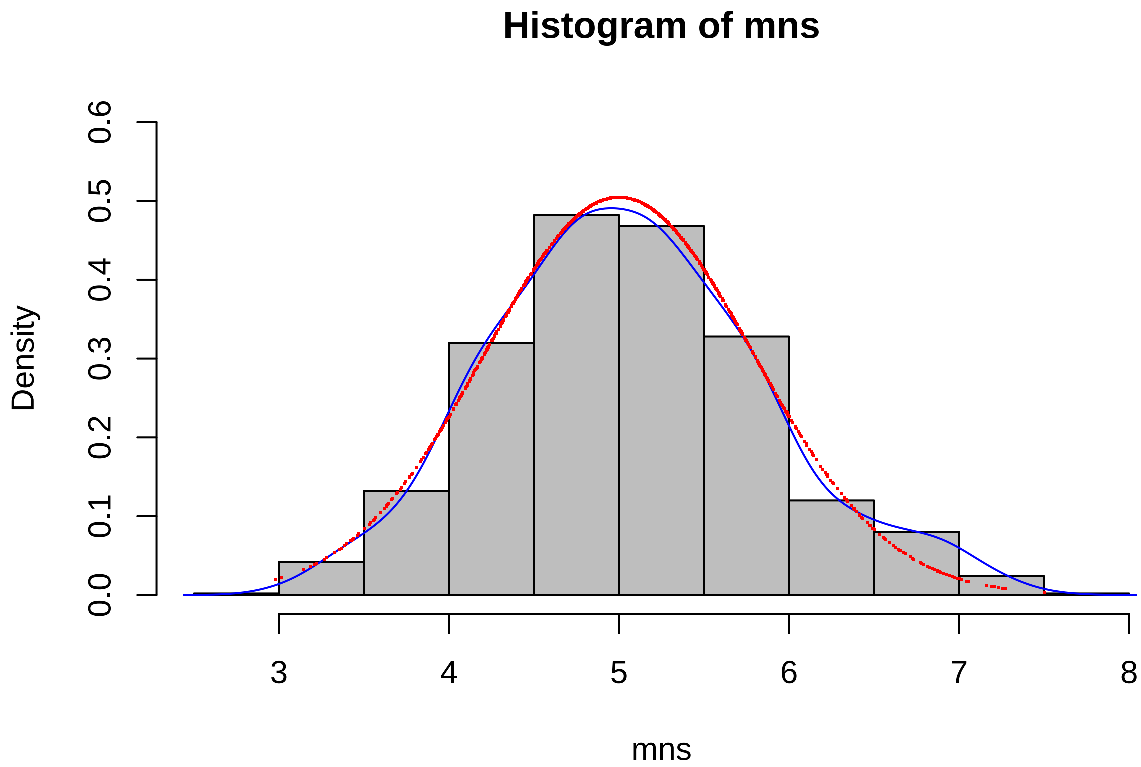
```
## [1] 0.625
```

```
actual_var   = var(mns)
actual_var
```

```
## [1] 0.6543703
```

Finally to see that the distribution is roughly normal we can plot a normal distribution over top of it and see that they are rather close indeed. The actual normal is plotted in red, the histogram and approximate density are plotted in grey and blue.

```
hist(mns,prob=TRUE,col='grey',ylim=c(0,.6))
lines(density(mns),col='blue')
points(mns,dnorm(mns,5,sqrt(.625)),col='red',pch='.')
```


Histogram of mns
Density
0.0
0.1
0.2
0.3
0.4
0.5
0.6
3
4
5
6
7
8
mns