Central Limit Theorem

Michael Houck

August 23, 2015

Summary

The Central Limit Theorem states that, under certain cirsumstances, taking a sufficient number of sample means from a random variable is itself a random variable which will be normally distributed around the original population mean. This analysis will show this theory in pracice via simulation using the exponential distribution.

Experiment

Set-up

First, we will set up the experiment variables: the random number seed, the number of draws, the number of simulation iterations and lambda.

```
set.seed(1234)
n <- 40
iterations <- 1000
lmb <- 0.2
```

Theoretical Mean

The theoretical mean of the exponential distribution is λ^{-1} and the theoretical variance is λ^{-2}

```
t.m <- lmb^-1
t.v <- lmb^-2

t.m
```

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

```
t.v
```

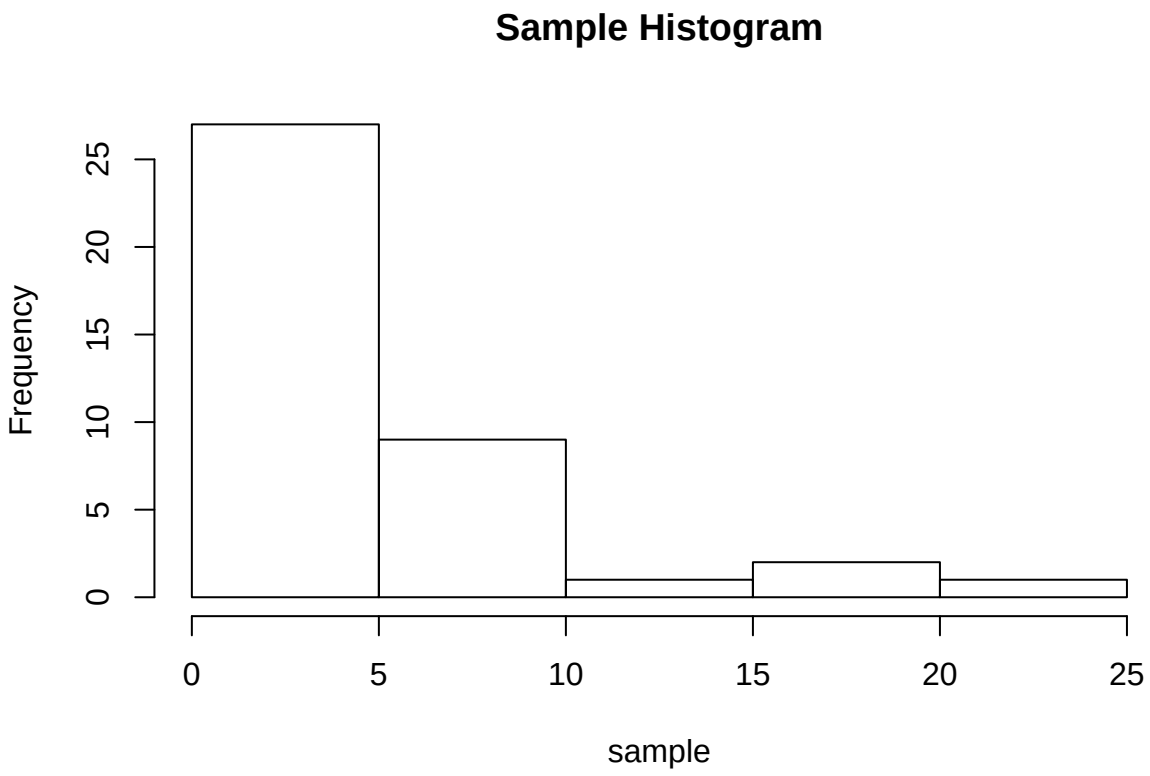
```
## [1] 25
```

Sample plot

If we plot a sample exponential distribution of n variables we see the structure of the distribution. We can also calculate the sample's mean and variance.

```
sample <- rexp(n,lmb)

hist(sample, main = "Sample Histogram")
```




```
s.m <- mean(sample)
s.v <- var(sample)

s.m
```

```
## [1] 4.969024
```

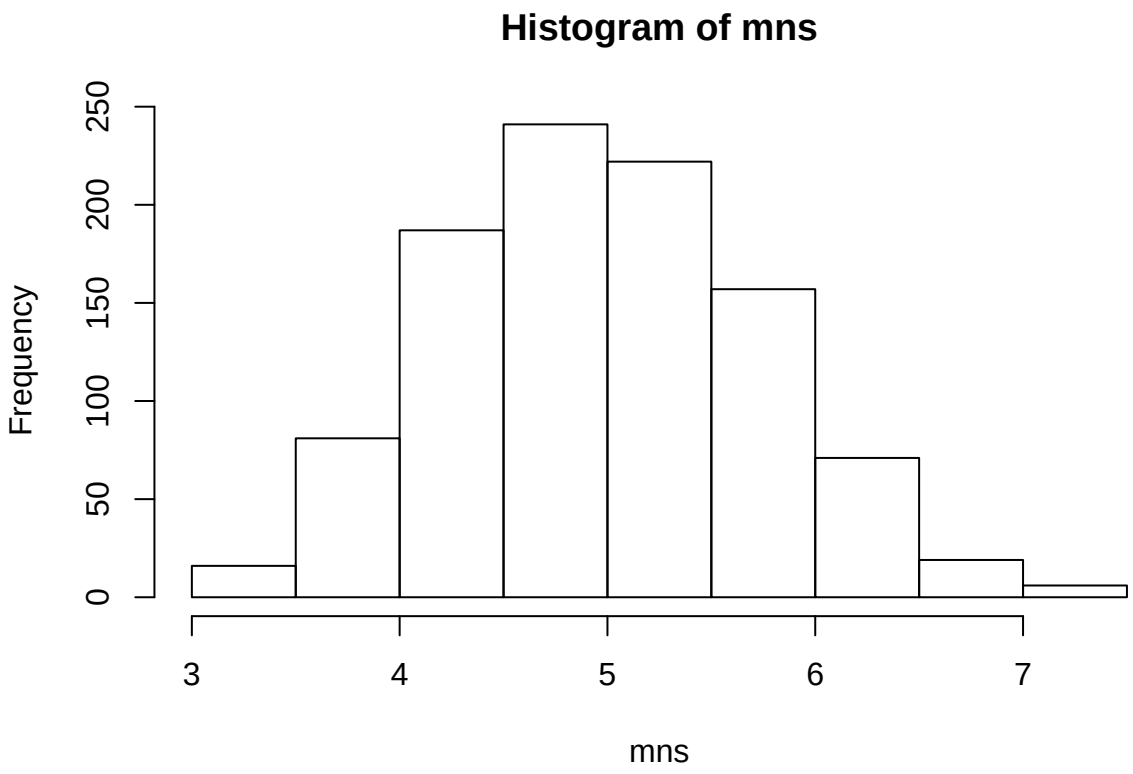
```
s.v
```

```
## [1] 25.99297
```

Simulation

When we run the simulation, we see the following:

```
mns = NULL
for (i in 1 : iterations) mns = c(mns, mean(rexp(n, lmb)))
hist(mns)
```




```
e.m <- mean(mns)
e.v <- var(mns)

e.m
```

```
## [1] 4.973641
```

```
e.v
```

```
## [1] 0.5710181
```

Comparisons

We can now compare the theoretical mean and variance with the mean and variance of one sample versus the mean and variance of the experimental samples.

Theoretical:

```
t.m
```

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

```
t.v
```

```
## [1] 25
```

One Sample:

```
s.m
```

```
## [1] 4.969024
```

```
s.v
```

```
## [1] 25.99297
```

Experiment:

```
e.m
```

```
## [1] 4.973641
```

```
e.v
```

```
## [1] 0.5710181
```

We can see that the theoretical mean and the one sample mean and variance are very similar. The experimental mean is even closer to the theoretical mean while the variance of the experiment is far less, as the distribution of the experiment is normal.