Statistical Inference Part 1 - Course Project

Clay Burns

June 18, 2015

Purpose

This report will explore the exponential distribution in R and compare it with the Cnetral Limit Theorem. This distribution can be simulated in R with rexp(n, lambda) where lambda is the rate paratmeter. The mean of the distrubtion is 1/lamdba and the stand deviation is also 1/lambda. The set lambda = .2 for all of the simulations. The exploration will cover a distribution of 40 exponentials and have a thousand simlultions.

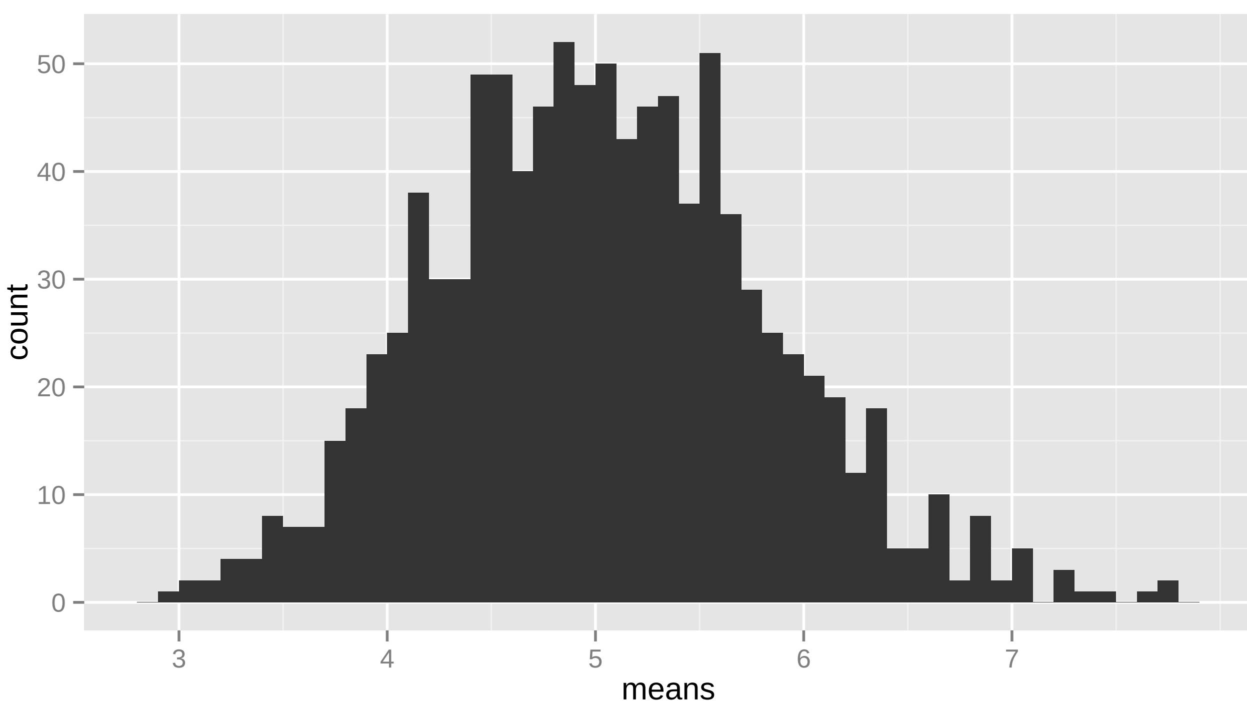
Simulations

```r
# load neccesary libraries
library(ggplot2)

# set constants n(40) = exponetial, lambda(.2) for rexp, number of tests = 1000
lambda <- 0.2
n <- 40
numberOfSimulations <- 1000

# set the seed to create reproducability
set.seed(22678979)

# run the test resulting in n(40) x numberOfSimulations(1000) matrix
exponentialDistributions <- matrix(data=rexp(n * numberOfSimulations, lambda), nrow=numberOfSimulations)
exponentialDistributionMeans <- data.frame(means=apply(exponentialDistributions, 1, mean))
```



Sample Mean versus Theoretical Mean

The expected mean μ of a exponential distribution of rate λ is

$\mu = \frac{1}{\lambda}$

```
mu <- 1/lambda
mu
```

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

Let $\bar{X}$ be the average sample mean of 1000 simulations of 40 randomly sampled exponential distributions.

```
meanOfMeans <- mean(exponentialDistributionMeans$means)
meanOfMeans
```

```
## [1] 5.035169
```

The two figures are very close.

Sample Variance versus Theoretical Variance

The expected standard deviation σ of a exponential distribution of rate λ is

$\sigma = \frac{1/\lambda}{\sqrt{n}}$

The e

```
sd <- 1/lambda/sqrt(n)
sd
```

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

The variance Var of standard deviation σ is

$Var = \sigma^2$

```
Var <- sd^2
Var
```

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

Let Var_x be the variance of the average sample mean of 1000 simulations of 40 randomly sampled exponential distribution, and σ_x the corresponding standard deviation.

```
sd_x <- sd(exponentialDistributionMeans$means)
sd_x
```

```
## [1] 0.7983913
```

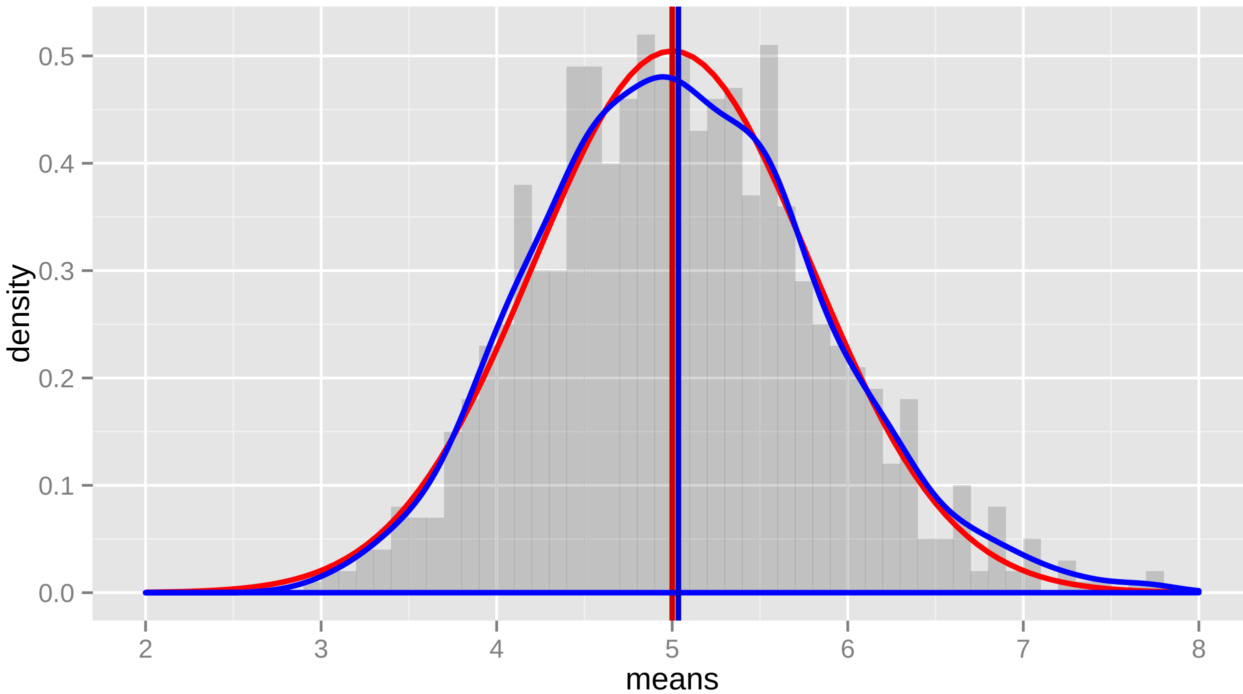
```
Var_x <- var(exponentialDistributionMeans$means)
Var_x
```

```
## [1] 0.6374287
```

Both figures are quite close but since variance is squared minor difference will likelyy be enhanced more.

Distribution

Time to compare the population means & standard deviation with a normal distribution of the expected values. Lines will be added for the calculated and expected means



The graph show just how true the central limit theorm is by how nicely the mean lines line up.