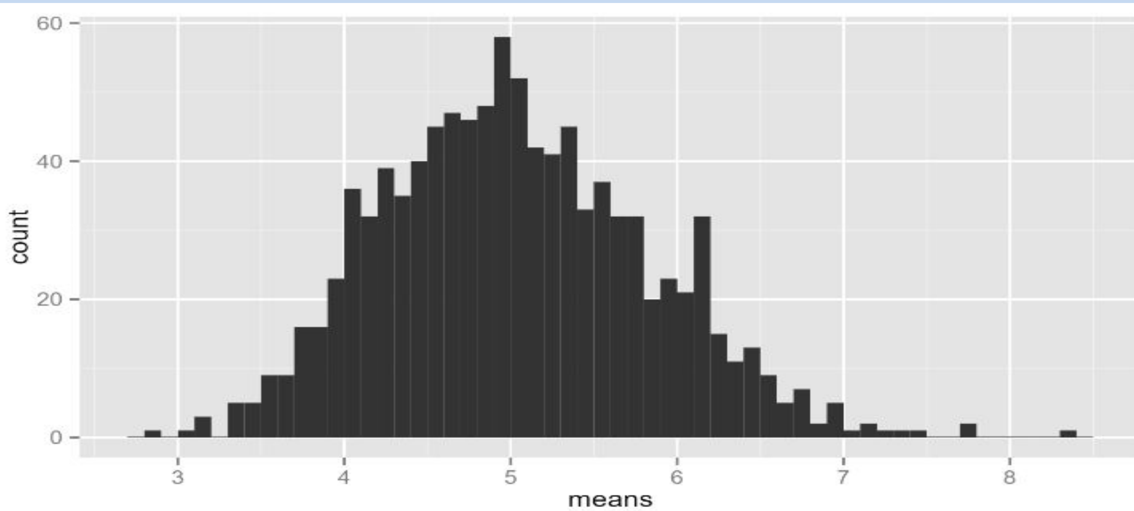
Statistical Inference Course Project 1

Overview

In this project I will investigate the exponential distribution in R and compare it with Central limit Theorem. The exponential distribution can be simulated in R with rexp(n,lambda), where lambda is the rate parameter. The mean of exponential distribution is 1/lambda and the Standard Deviation is also 1/lambda. I will set lambda=0.2 for all the simulations.I will investigate the distribution of averages of 40 exponentials. Note that 1000 simulations will be required.

Simulations

```
# load necessary libraries

library(ggplot2)

set constants

lambda<-0.2  #lambda for rexp

n<-40 #number of exponentials

numberOfSimulations<-1000 #number of tests

# set seed to create reproducibility

set.seed(11081979)

# run the test resulting in n x numberOfSimulations matrix

exponentialDistributions <- matrix(data=rexp(n*  numberOfSimulations, lambda), nrow=
numberOfSimulations exponentialDistributionsMeans <-
data.frame(means=apply(exponentialDistributions, 1 , mean) Sample
```



Sample Mean versus Theoretical Mean

The expected mean μ of an exponential distribution λ is

$\mu = 1/\lambda$

```
mu<- 1/lambda

mu
```

[1] 5

Let Ø be the average sample mean of 1000 simulations of 40 randomly sampled exponential distributions.

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

meanOfMeans
```

[1] 5 .027126

As seen above the expected mean and the average sample mean are very close.

Sample Mean versus Theoretical Mean

The expected Standard Deviation σ of an exponential distribution of rate λ is

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

The e

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

sd
```

[1] 0.7905694

Then variance Var of the standard deviation σ is

$Var = \sigma^2$

```
Var <- sd^2

Var
```

##[1] 0.625

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

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

sd_x
```

[1] 0.8020334

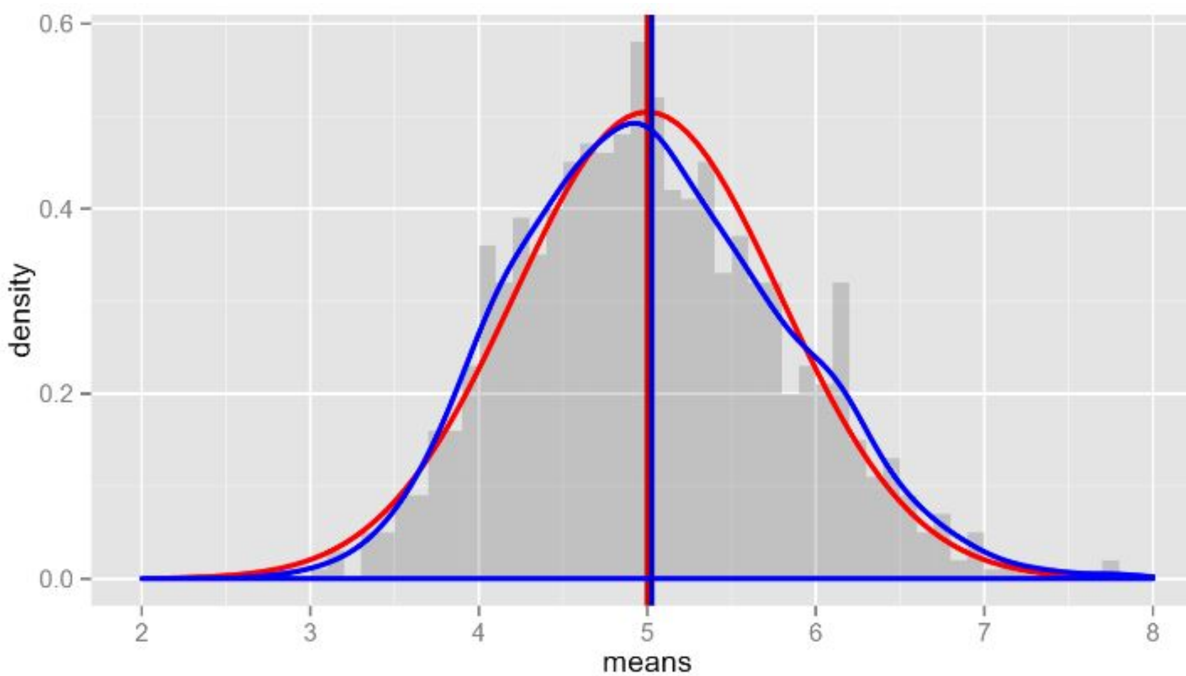
```
Var_x<- var(exponentialDistributionMeans$means)

Var_x
```

[1] 0.6432577

As evident above, the Standard deviations are very close. Since variance is the square of the standard deviations, minor differences will be enhanced, but are still pretty close.

Distribution

Comparing the population means and standard deviation with a normal distribution of the expected values . Added lines for calculated and expected means.



As you can see from the graph, the calculated distribution of means of random sampled exponential distributions , overlaps quite nice with the normal distribution with the expected values based on the given lambda.