# Exponential distribution Compare with the Central Limit Theorem

# Overview

In this project you will investigate the exponential distribution in R and compare it with the 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. Set lambda = 0.2 for all of the simulations. You will investigate the distribution of averages of 40 exponentials. Note that you will need to do a thousand simulations.

# Introduction

Illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponentials. You should

1. Show the sample mean and compare it to the theoretical mean of the distribution.
2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.
3. Show that the distribution is approximately normal.

In point 3, focus on the difference between the distribution of a large collection of random exponentials and the distribution of a large collection of averages of 40 exponentials.

As a motivating example, compare the distribution of 1000 random uniforms

hist(runif(1000))

and the distribution of 1000 averages of 40 random uniforms

mns = NULL  
for (i in 1 : 1000) mns = c(mns, mean(runif(40)))  
hist(mns)

This distribution looks far more Gaussian than the original uniform distribution!

# Output

* Sample mean vs theoretical mean of the distribution.

We generate a series of 1000 simulation dataset for comparison to the theory. Each iteration contains 40 observation and the exponential distribution is set using 'rexp' function with lambda = 0.2

Below is the variables and its value :-

lambda = 0.2  
n = 40  
nosim = 1000

Now we perform the code below to simulate , collect the data and plot it.

rexp() - will generate 40 times when the rate is 0.2 and get the mean value from it.

exponen\_simulation <- function(n, lambda)  
{  
mean(rexp(n,lambda))  
}  
  
simData <- data.frame(ncol=2,nrow=1000)  
names(simData) <- c("Index","Mean")  
for (i in 1:nosim)  
{  
simData[i,1] <- i  
simData[i,2] <- exponen\_simulation(n,lambda)  
}

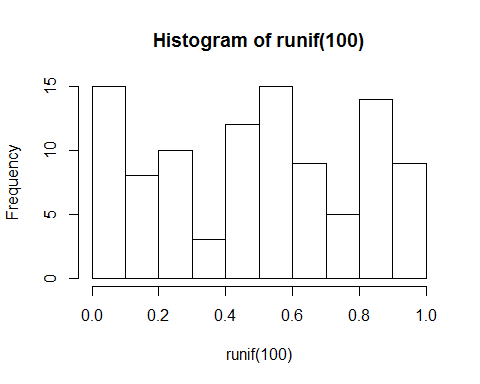
The mean of the sample dataset is :-

sample\_mean <- mean(simData$Mean)  
 sample\_mean

## [1] 5.050252

* Variance variable vs theoretical variance of the distribution.

hist(runif(100))



* Distribution is approximately normal.

hist(runif(10))

