Statistical Inference Project Part 1

Andres P

29/01/2019

Overview

This project investigates the exponential distribution of R and applies The Central Limit Theorum to a thousand mean simulations.

Exponential distribution is simulated in R using rexp(n, lambda). Lambda is the rate perameter. The mean of exponential distribution is 1/lambda. The standard deviation is 1/lambda.

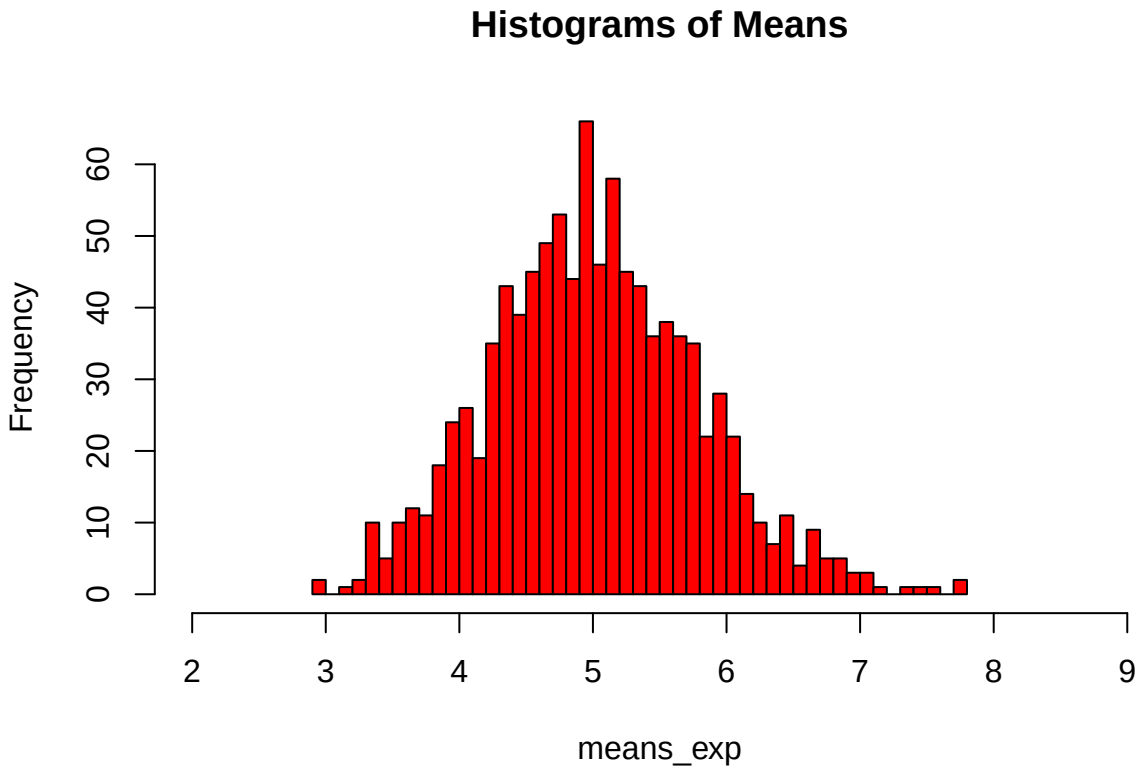
Instructions provided are as follows.

lambda = .2 for all simulations Investigate the distribution of averages of 40 exponentials Do 1,000 simulations

Simulations Explanation

This code sets the parameters as an outline

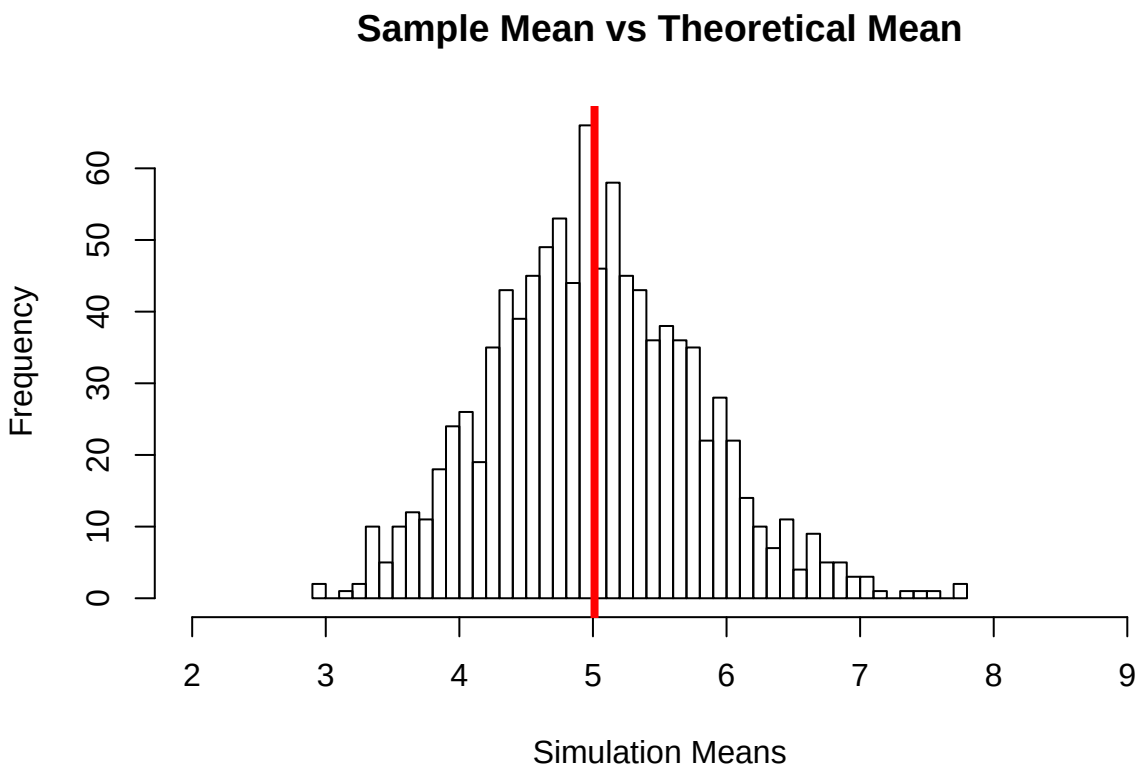
```
# set seed for reproducability
set.seed(123)

# set sampling values as mandated in the instructions

#lambda
lambda <- 0.2

# number of exponentials
n <- 40

# number of simulations
run_sim <- 1000

# Run simulations

run_sim<- replicate(run_sim,rexp(n,lambda))

# Calculate the mean of the exponential simulations

means_exp<- apply(run_sim,2,mean)

#Histogram of the means

hist(means_exp, breaks=40, xlim = c(2,9), main="Histograms of Means", col = "red")
```



Sample Mean vs Theoretical Mean

The exponential distribution mean is equal to 1/lambda. For this simulation, lambda is 0.2. The theoretical mean is 5.

```
# plot histogram of the sample means
hist(means_exp, main = "Sample Mean vs Theoretical Mean", xlim =  c(2,9), breaks = 40, xlab = "Simulati

# plot vertical blue line at mean of samples
abline(v = mean(means_exp), lwd = "4", col = "red")
```




```
# calculate the sample mean
mean(means_exp)
```

```
## [1] 5.011911
```

```
## [1] 5.011911
```

The sample mean of 5.011911 is close to the theoretical mean of 5

Sample Variance vs Theoretical Variance

The standard deviation of the exponential distribution is (1/lambda) / sqrt(n).

```
# theoretical variance vs simulated variance
print(paste ("Theoretical variance is: ", round( (1/lambda)^2/n, 3)))
```

```
## [1] "Theoretical variance is:  0.625"
```

```
print(paste("Actual variance is: ", round( var(means_exp),3)))
```

```
## [1] "Actual variance is:  0.6"
```

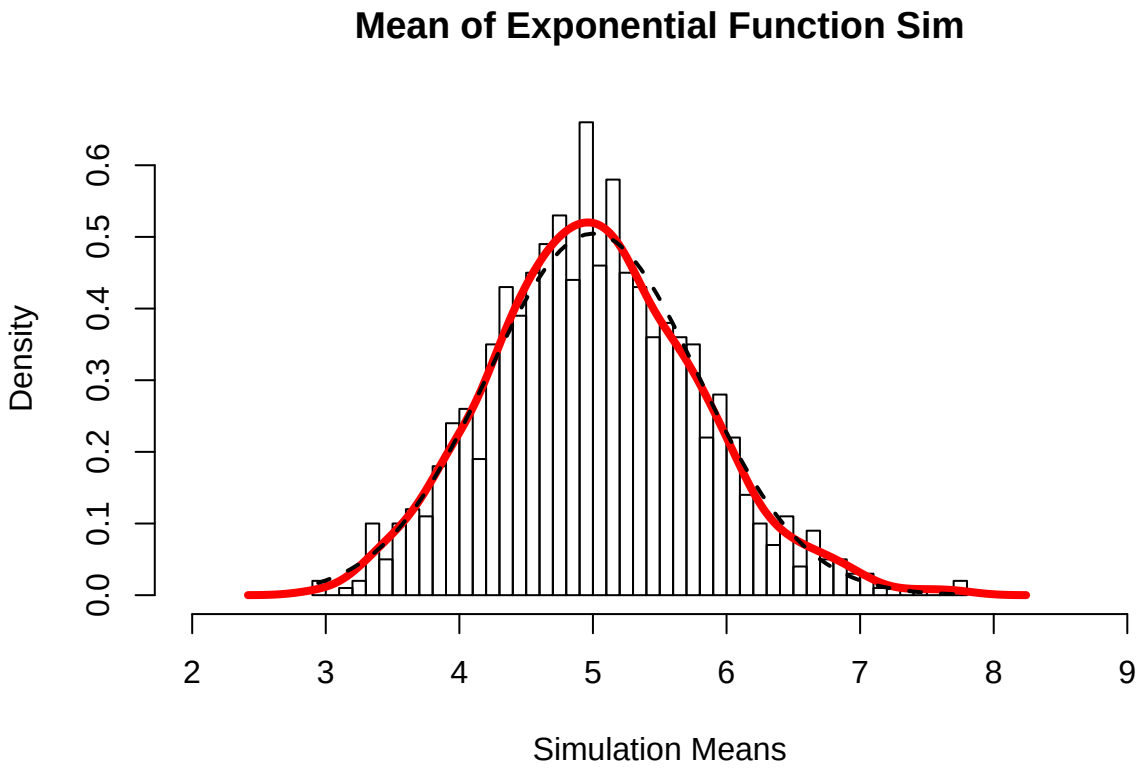
The actual varience is slightly lower than the theoretical variance.

Distribution

The histogram below shows whether the exponential distribution is approximately normal. Due to the Central Limit Theorem, the means of the sample simulations should follow a normal distribution.

```
# Histogram with distribution curve
hist(means_exp, prob=TRUE, main = "Mean of Exponential Function Sim", breaks = 40, xlim = c(2,9), xlab =
```

```r
lines(density(means_exp), lwd=4, col="red")

# Normal distribution line
x <- seq(min(means_exp), max(means_exp), length = 2*n)
y <- dnorm(x, mean = 1/lambda, sd = sqrt(((1/lambda)/sqrt(n))^2))
lines(x,y, pch = 20, lwd = 2, lty = 2)
```



The histogram above shows the calculated distribution of means of the simulated exponentials distributions is a normal distribution because of the Central Limit Theorum. If the number of samples increases the distribution would be closer to the standard normal distribution.