Title: “Statistical Inference Course Project Part 1”

author:“Sohair Zaki” **date:**“May 12, 2017”

Part 1: Simulation Exercise In this project we 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. We will investigate the distribution of averages of 40 exponentials. We need to do a thousand simulations. pseodocode Illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponentials. You should

–Show the sample mean and compare it to the theoretical mean of the distribution. –Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution. –Show that the distribution is approximately normal. — start r code

```
# 1. Show the sample mean and compare it to the theoretical mean of the distribution.
    # Set seed
    set.seed(2)

    # Set lambda per value provided
    lambda <- 0.2

    # Simulations 1000 simulation
    sim <- 1000

    # Samples we are given 40 exponentials
    n <- 40

    #exponential distribution can be simulated in r using rexp
    # we need to replicate that 1000 times per requirements 1000 simulations

    ExpSimulation <-  replicate(sim, rexp(n,lambda))

    #take the mean
    ExpMean <- colMeans (ExpSimulation)

    # simulated mean
    SimuMean <- mean (ExpMean)

    # let us display it
    SimuMean
```

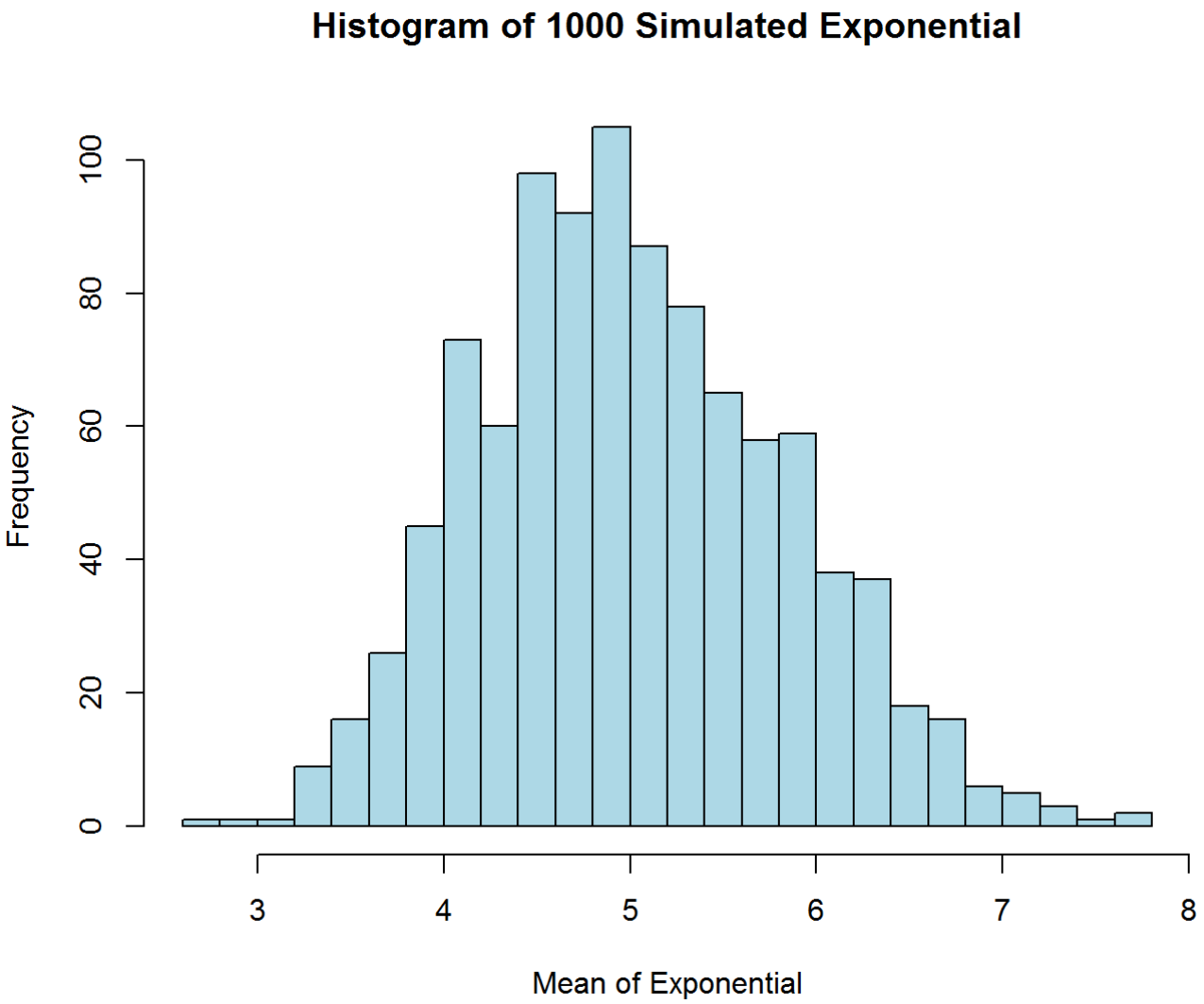
```
## [1] 5.016356
```

```
    #note it is equal to 5.016356

    # Therorical mean is 1/lambda
    TheroticalMean <- 1/0.2
    #let us display it
    TheroticalMean
```

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

```
    # note it is equal to 5
 # conclusion : The simulated mean is 5.01 and  theoretical mean is 5.
```

```
# plot size is 7 inches wide and 6 inches high
 # let us draw a histogram for the mean of the simulated exponential ==ExpMean
# xlab=label of x axis is Mean of Exponential
# main= title of the graph Histogram of 1000 simualtions
# col= color of the bars
hist(ExpMean, breaks = 20, xlab = "Mean of Exponential",
     main = "Histogram of 1000 Simulated Exponential",
     col = "lightblue")
```



```
#Show how variable the sample is (via variance) and compare it to the theoretical variance of th
e distribution.
# get the standard deviation of ExpMean using sd()

SimuSTD <- sd(ExpMean)
# get its variance

SimuVAR <- SimuSTD^2

# let us display it
SimuVAR
```

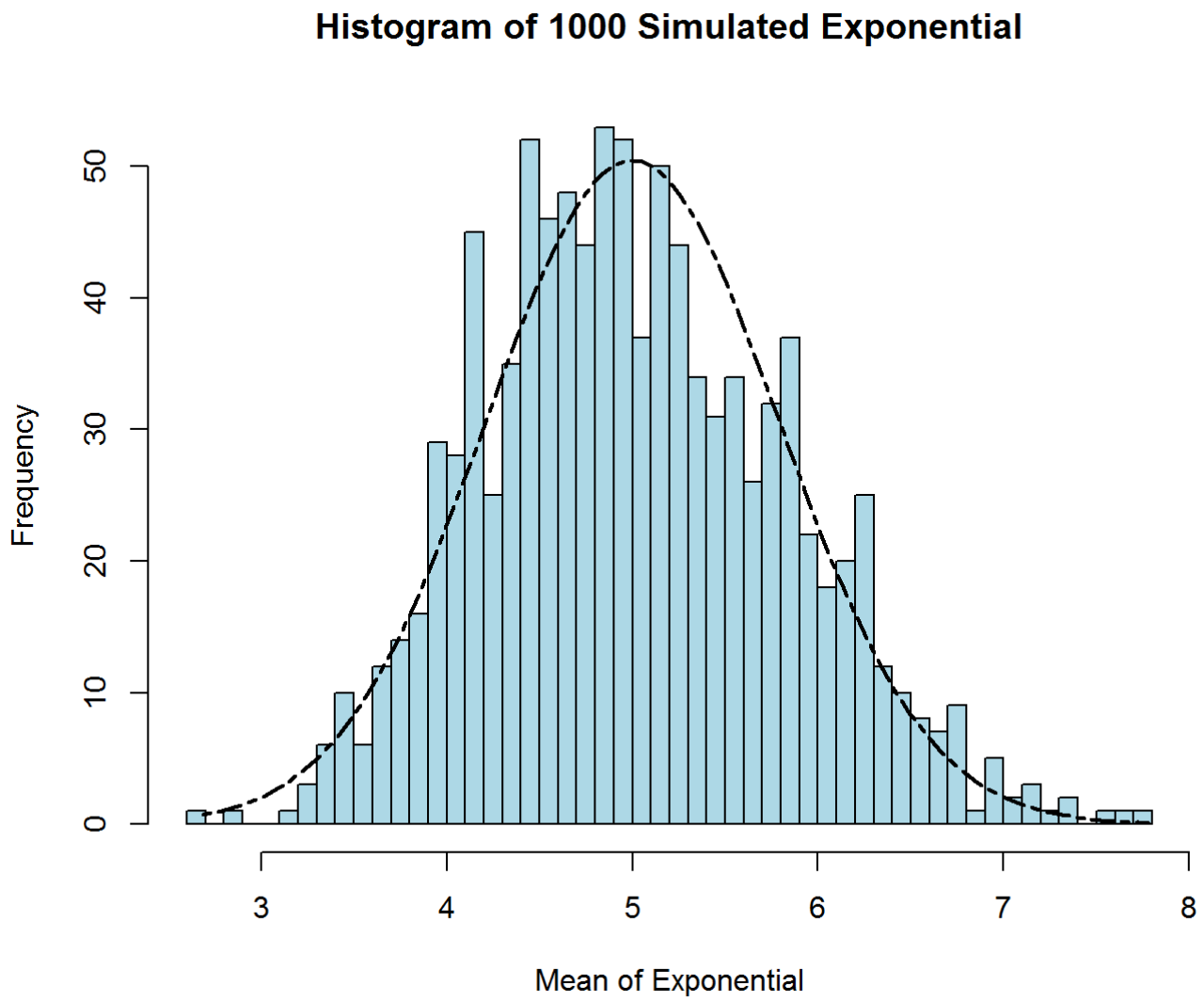
```
## [1] 0.6691305
```

```
#Note the result is 0.6691305

# Theoretical Standard Deviation is 1/lambda/sqrt (n)
TheroticalSTD <- (1/lambda)/sqrt(n)
# Theoretical variance
TheroticalVAR <- TheroticalSTD^2

# let us display it
TheroticalVAR
```

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

```
# the result is 0.625

# conclusion The simulated variance is 0.6691 cand theoretical variance of 0.625.

#3. Show that the distribution is approximately normal
# Plot Histogram
```

```
# plot size is 7 inches wide and 6 inches high
 # let us draw a histogram for the mean of the simulated exponential ==ExpMean
# xlab=label of x axis is Mean of Exponential
# main= title of the graph Histogram of 1000 simualtions
# col= color of the bars
hist(ExpMean, breaks = 40, xlab = "Mean of Exponential",
     main = "Histogram of 1000 Simulated Exponential",
     col = "lightblue")

# Add the Theoretical Normal Distribution Line
# generate a seq
x  <- seq(min(ExpMean), max(ExpMean), length = 100)

y  <- dnorm(x , mean = 1/lambda, sd = 1/lambda/sqrt(n))
# draw lines with line type 6


lines(x , y *100, lty=6, lwd=2)
```



Conclusion :the distribution is closely matched with a normal distribution.