Exponential Distribution Simulation & Analysis

This report will investigate the exponential distribution in R and compare it with the Central Limit Theorem (CLT). The basis of the investigation will be 1000 repeated simulations of 40 exponentially distributed random variables. The sample mean, sample variance, and distribution of the averages of these 1000 simulations will be compared to that predicted by the CLT.

We begin by specifying arbitrarily that lambda (rate) is 0.2 and, as previously noted, simulate 40 exponentially distributed random variables 1000 times.

```
set.seed(8)
nosim <- 1000
n <- 40
rate <- 0.2
sim_means <- replicate(nosim, mean(rexp(n, rate)))
```
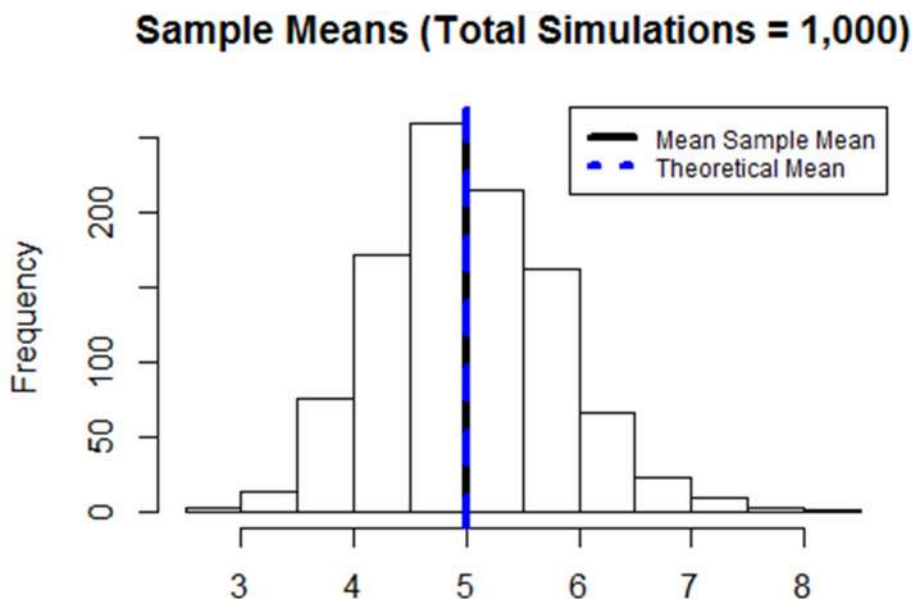
1. Show the sample mean and compare it to the theoretical mean of the distribution.

Find the sample mean:

```
mean(sim_means)
```

```
## [1] 5.006442
```

The sample mean of 5.0064417 is very close to the theoretical mean of 1/rate or 5.

```
hist(sim_means, main="Sample Means (Total Simulations = 1,000)",xlab="")
abline(v = mean(sim_means),lwd=4)
abline(v = 1/rate, col = "blue",lty=2,lwd=4)
legend("topright",col = c("black", "blue"),lty = c(1,3), lwd=4,
       legend = c("Mean Sample Mean","Theoretical Mean"), cex=.75
       )
```

2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.

Find the sample variance

```
var(sim_means)
```

```
## [1] 0.6028838
```

The sample variance of 0.6028838 is also very close to the theoretical variance predicted by the CLT or 0.625.

```
x = 30:50
y = (1/rate/sqrt(x))^2
plot(40,var(sim_means), main=expression(paste("Theoretical variance = ",
                                        (over(frac(1,lambda),sqrt(n)))^2)),
      xlab="Sample sizes", ylab="Theoretical Variance", pch=15
      )
lines(x,y,col="blue",lwd=3)
text(40,var(sim_means),"Sample Variance", pos=2)
text(50,(1/rate/sqrt(50))^2,"Theoretical Variance", pos=1)
```
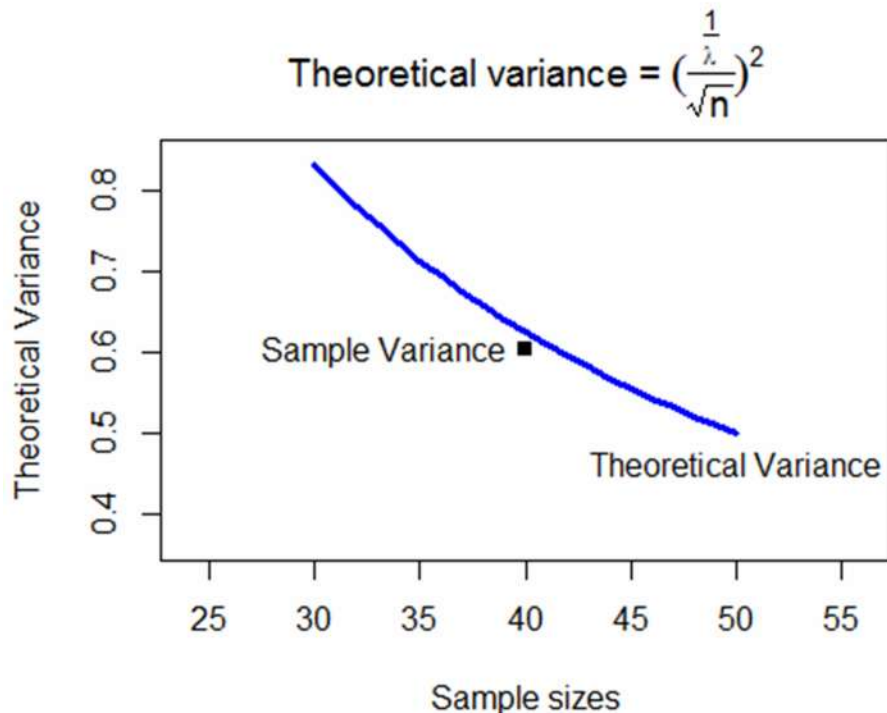

3. Show that the distribution is approximately normal.

The means of the simulated data are approximately normally distributed as predicted by the CLT. Below is histogram of the data with a normal PDF with the mean and standard deviation equal to the theoretical mean and standard error laid on top. It is clear the distributions are similar.

```
hist(sim_means, freq = FALSE,
      main="Sample Means (Total Simulations = 1,000)"
      , xlab=""
      )
lines(density(sim_means), lwd=2)
lines(seq(2.0, 8.3, by = 0.1),
      dnorm(seq(2.0, 8.3, by = 0.1), mean = 1/rate, sd = 1/rate/sqrt(n)),
      col = "blue", lwd=2
```

```
        )
legend("topright",col = c("black", "blue"), lwd=2,
        legend = c("Sample PDF","Normal PDF"), cex=.75
        )
```

Sample Means (Total Simulations = 1,000)