

Investigation of the exponential distribution in R and comparison with the CLT

Yanal Kashou

Contents

Overview of Assignment.....	1
Simulation.....	1
Sample Mean vs Theoretical Mean.....	1
Sample Variance vs Theoretical Variance	2
Conclusion	3

Overview of Assignment

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.

Simulation

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

```
# Load necessary libraries
library(ggplot2)

# We begin by setting values of lambda, samples, simulations and finally seed for reproducibility
lambda <- 0.2
n <- 40
samples <- 1000
set.seed(0)
```

Sample Mean vs Theoretical Mean

```
# create the exponential sample
exp_samp <- replicate(samples, rexp(n, lambda))
exp_mean <- colMeans(exp_samp)

# Sample mean
```

```
mean_samp <- mean(exp_mean)
mean_samp

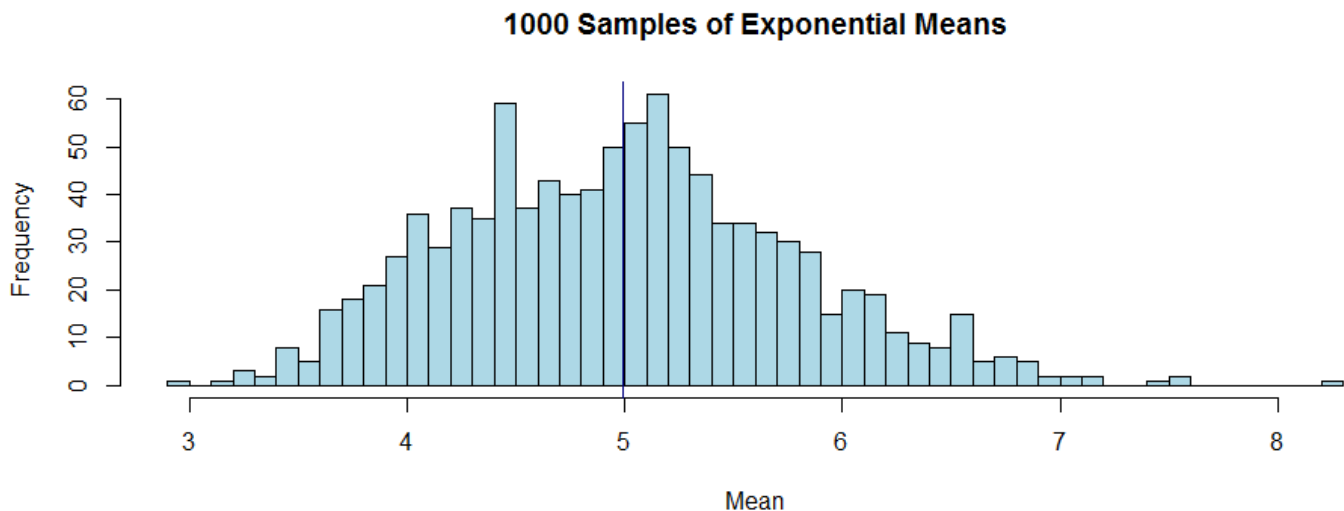
## [1] 4.989678

# Theoretical Mean
mean_theo <- 1/lambda
mean_theo

## [1] 5
```

Our sample mean is ~ 4.989 while the theoretical mean is $= 5$

```
# Plot
hist(exp_mean, breaks = 49, xlab = "Mean", main = "1000 Samples of Exponential Means",
     col = "light blue")
abline(v = mean_samp, col = "navy")
```



Sample Variance vs Theoretical Variance

```
# Variance
var(exp_mean)

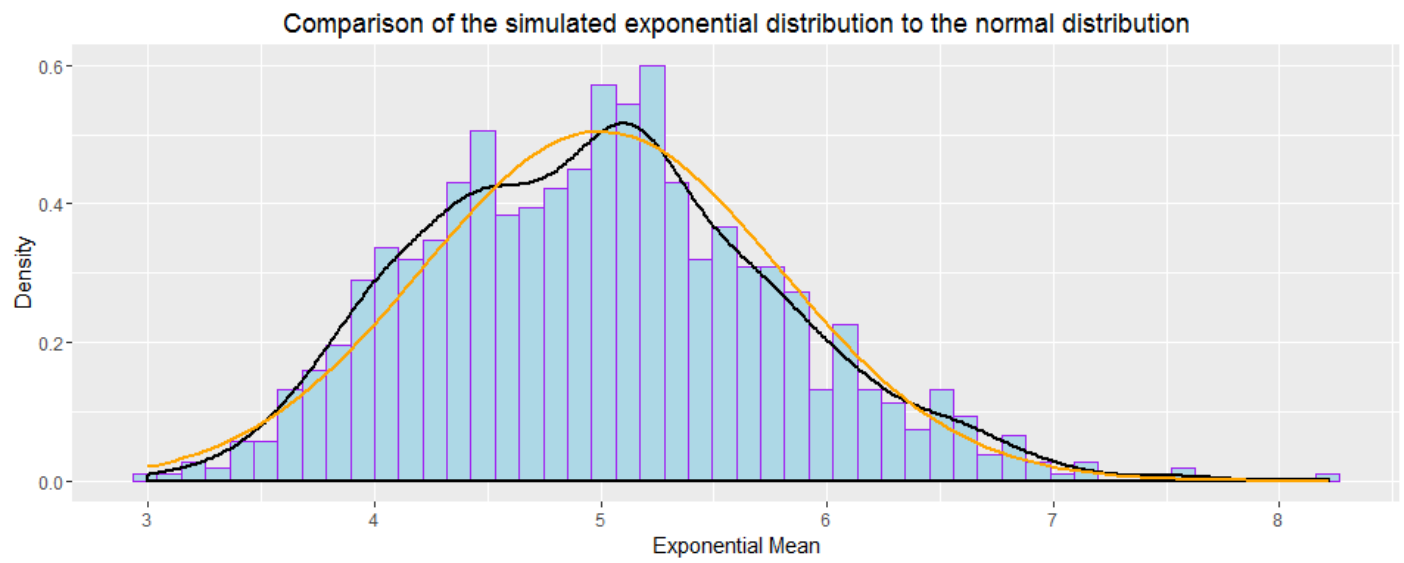
## [1] 0.6181582

(sig2 <- (1/lambda^2)/n)

## [1] 0.625
```

Our sample variance is ~ 0.618 while the theoretical variance is $= 0.625$

```
#### Distribution Final plot
ggplot(data.frame(exp_mean), aes(x = exp_mean)) + geom_histogram(aes(y = ..density..),
  colour = "purple", fill = "light blue", bins = 50) + geom_density(colour = "black",
  size = 1) + stat_function(fun = dnorm, args = list(mean = 5, sd = sqrt(sig2)),
  colour = "orange", size = 1) + xlab("Exponential Mean") + ylab("Density") +
ggtitle("Comparison of the simulated exponential distribution to the normal
distribution")
```



Conclusion

According to the plot, our simulation closely matches the normal distribution