

Simulation Exercise

Statistical Inference Course Project: Part 1

Harkishan Grewal

11/05/2020

Overview

In this project, I 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. I will investigate the distribution of averages of 40 exponentials.

Simulations

Simulate the distribution of 1000 averages of 40 random exponentials.

```
# set seed for reproducibility
set.seed(1000)

# number of exponentials
n = 40

# Set lambda = 0.2 for all of the simulations
lambda = 0.2

# number of simulations
samples <- 1000

# do 1000 random simulations
mns = NULL

for (i in 1 : samples) {
  mns = c(mns, mean(rexp(n, lambda)))
}
```

Sample Mean versus Theoretical Mean

```
# means of simulations
sample_mean <- mean(mns)

# theoretical mean
theoretical_mean <- 1 / lambda

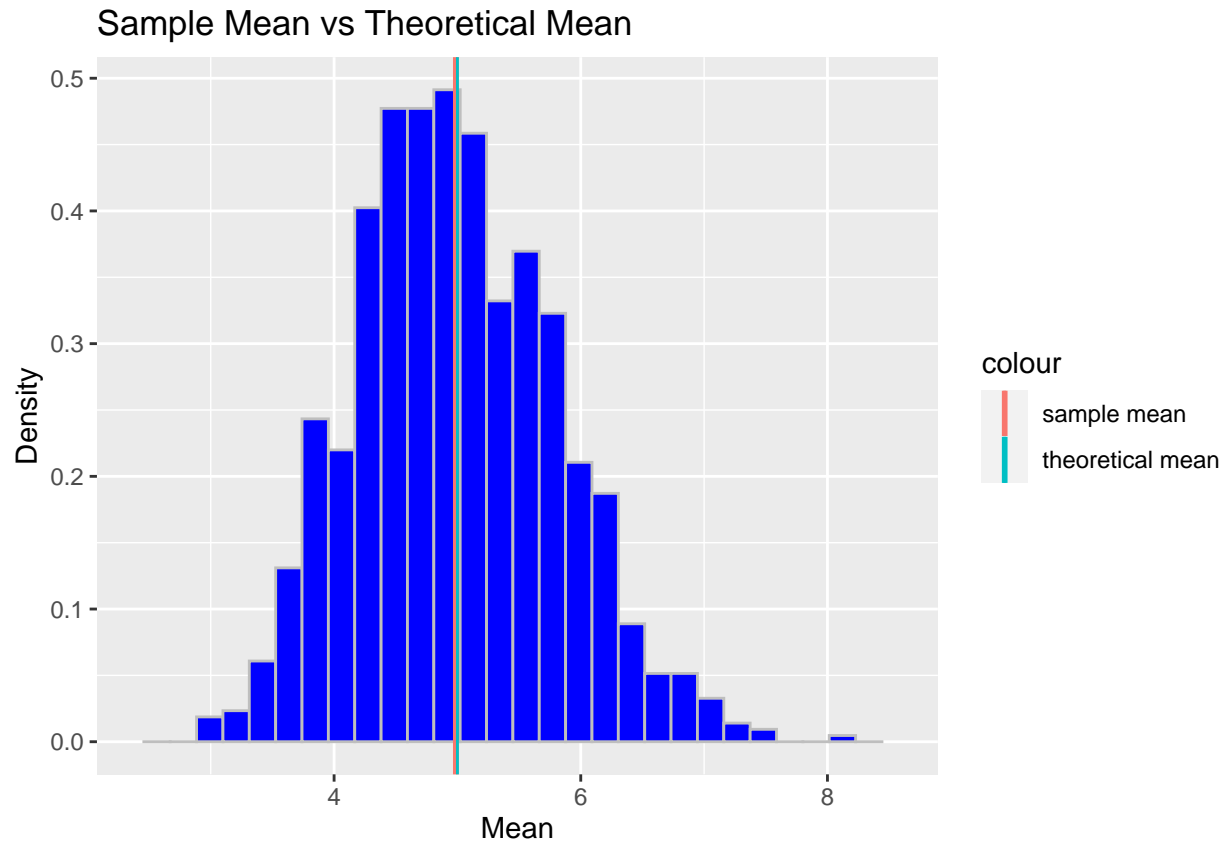
df <- data.frame(mns)

g <- ggplot(df, aes(x = mns))
```

```

g <- g + geom_histogram(aes(y = ..density..), fill = "blue", color = "grey")
g <- g + xlim(range(density(mns)$x))
g2 <- g + geom_vline(
  aes(xintercept = sample_mean, color="sample mean"),
  lwd = 1)
g2 <- g2 + geom_vline(
  aes(xintercept = theoretical_mean, color="theoretical mean"))
g2 <- g2 + labs(
  title = "Sample Mean vs Theoretical Mean", x = "Mean", y = "Density")
g2

```



```

# Print Sample Mean and Theoretical Mean
c(sample_mean, theoretical_mean)

```

```
## [1] 4.986963 5.000000
```

The sample mean is 4.9869634 which is close to the theoretical mean of 5.

Sample Variance versus Theoretical Variance

```

# variance of simulations
sample_var <- var(mns)

# theoretical variance
theoretical_var <- (1 / lambda)^2 / n

```

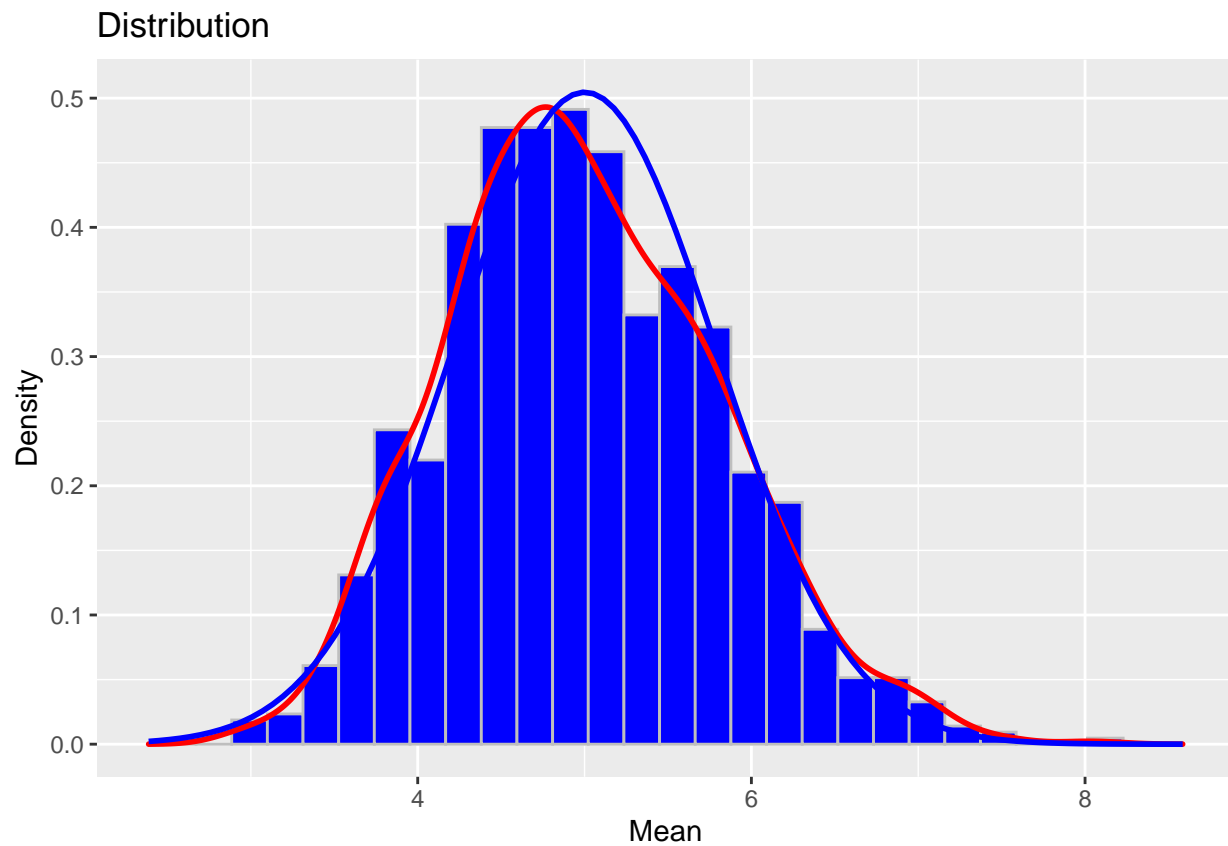
```
# Print Sample Variance and Theoretical Variance
c(sample_var, theoretical_var)
```

```
## [1] 0.654343 0.625000
```

The sample variance is 0.654343 which is close to the theoretical variance of 0.625.

Distribution

```
# theoretical standard deviation
theoretical_sd <- 1 / (lambda * sqrt(n))
g3 <- g + geom_density(aes(y = ..density..), color = "red", lwd = 1.0)
g3 <- g3 + stat_function(
  fun = dnorm, args = list(mean = theoretical_mean, sd = theoretical_sd),
  color = "blue", lwd = 1.0)
g3 <- g3 + labs(title = "Distribution", x = "Mean", y = "Density")
g3
```



The distribution (red line) is approximately normal (blue line).