simlulation-exercise

Overview

In this project, I will investigate the exponential distribution in R and compare it with the Central Limit Theorem. I will compare the sample mean and variance with the theoretical mean and variance. I will plot the distribution and explain how it is approximately normal.

Simulations

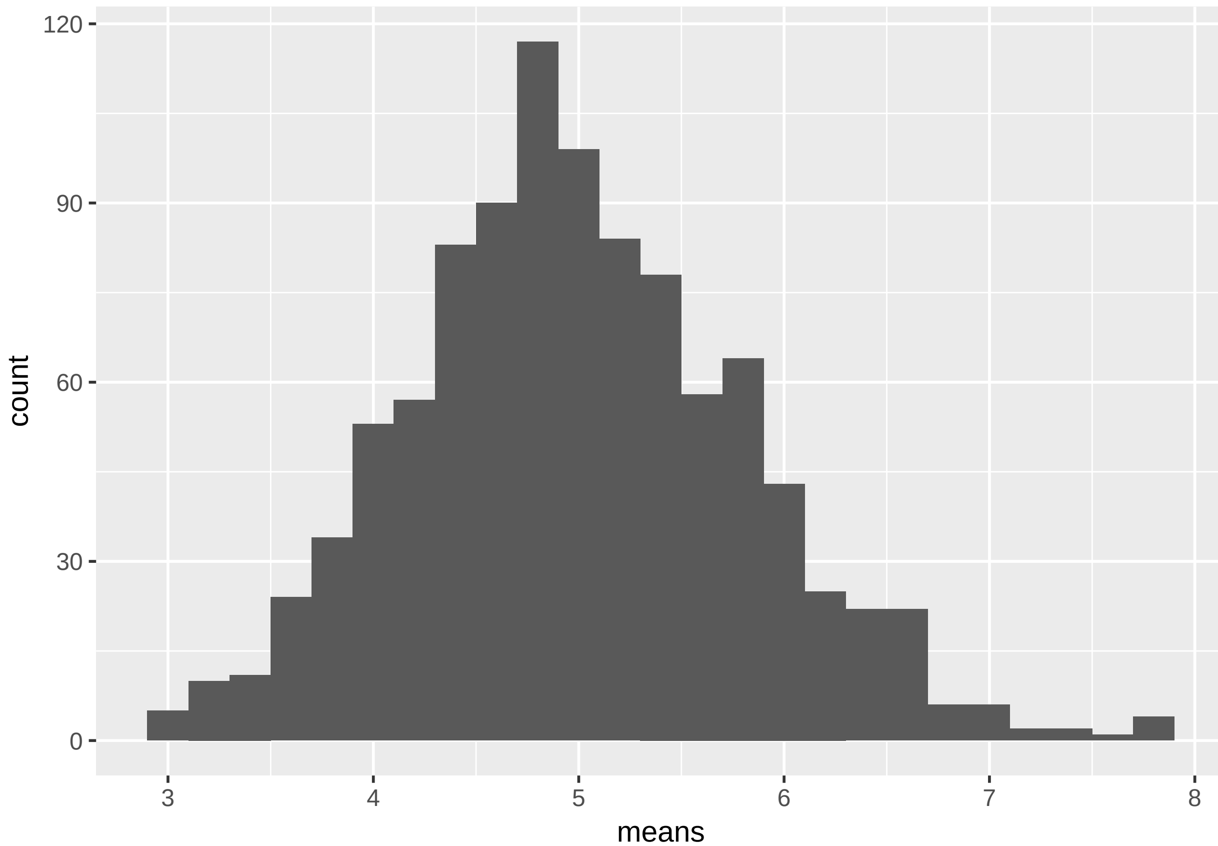
```
library(ggplot2)

# Set constants provided in the assignment instruction
lambda <- 0.2
number_of_exponetials <- 40
number_of_simulations <- 1000

set.seed(0)

# Run the simulation and calculate the means
exponential_matrix = matrix(data=rexp(number_of_exponetials * number_of_simulations, lambda), nrow=numbe
exponential_means = data.frame(means=apply(exponential_matrix, 1, mean))

# plot the means for comparison with the Central Limit Theorem
ggplot(data = exponential_means, aes(x = means)) + geom_histogram(binwidth=0.2)
```



Compare the sample mean with the theoretical mean

```
theoretical_mean <- 1/lambda
theoretical_mean
```

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

```
sample_mean <- mean(exponential_means$means)
sample_mean
```

```
## [1] 4.989678
```

They are pretty much the same.

Compare the sample variance with the theoretical variance

```
theoretical_variance <- (1/lambda/sqrt(number_of_exponetials))^2
theoretical_variance
```

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

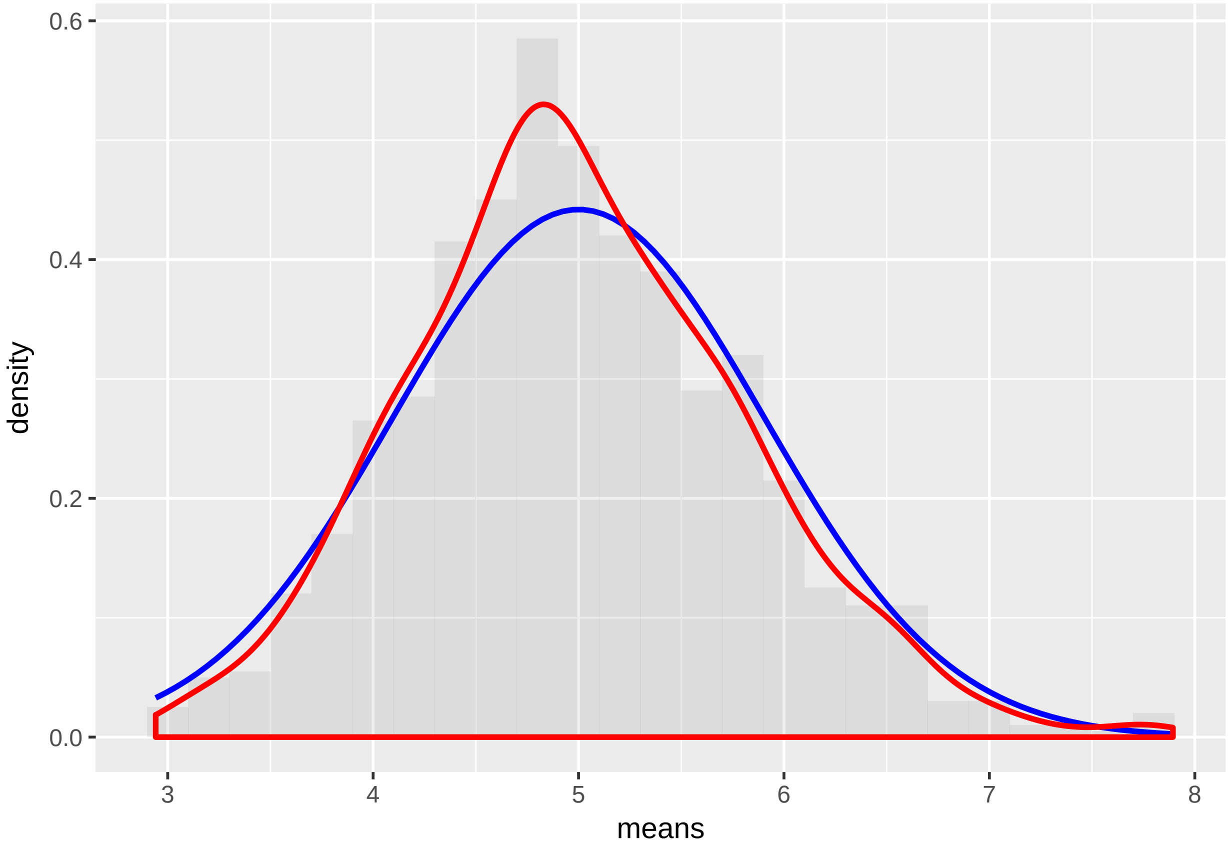
```
theoretical_variance <- sd(exponential_means$means)
theoretical_variance
```

```
## [1] 0.8147837
```

The difference here is larger than the one regarding the mean, but they are still pretty close.

Compare the sample distribution with the one of the Central Limit Theorem

```
ggplot(data = exponential_means, aes(x = means)) +
  geom_histogram(aes(y=..density..), alpha=0.1, binwidth=0.2) +
  stat_function(fun = dnorm, args = list(mean = theoretical_mean, sd = sqrt(theoretical_variance)), colo
  geom_density(colour="red", size=1)
```



The two graphs overlap quite well with each other.