Project of simulation exercise

Zhibin Guan

Overview

In this we will investigte 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 averates of 40 exponentials. We will do a thousand simulations

Demands

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

1. Show the sameple mean and compare it to the theoretical mean of the distribution.
2. Show the variable the sample is (via variance) and compare it to the theoretical variance of the distribution.
3. Show that the distribution is approximately normal.

Answers

Pre-processing for the project

Create random numbers which respect to the exponential distribution

```
set.seed(10)
# exponent -- n is 40
n <- 40
lambda <- .2

df_sim <- NULL
for (i in 1:1000) {
    df_sim <- c(df_sim, mean(rexp(n, lambda)))
}
```

```
# theoretical mean
tm <- 1/lambda
# theoretical variance
tv <- ( 1/(sqrt(n) * lambda) )^2
```

```
sm <- mean(df_sim)
sv <- round(var(df_sim), 2)
```

Answer of demand 1

Show sample mean (sm), and compare it to the theoretical mean (tm):

```
sm
```

```
## [1] 5.04506
```

```
tm
```

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

```
abs(sm - tm)
```

```
## [1] 0.04505959
```

Answer of demand 2

Show sample variable (sv), and compare it to the theoretical variable (tv):

```
sv
```

```
## [1] 0.64
```

```
tv
```

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

```
abs(sv - tv)
```

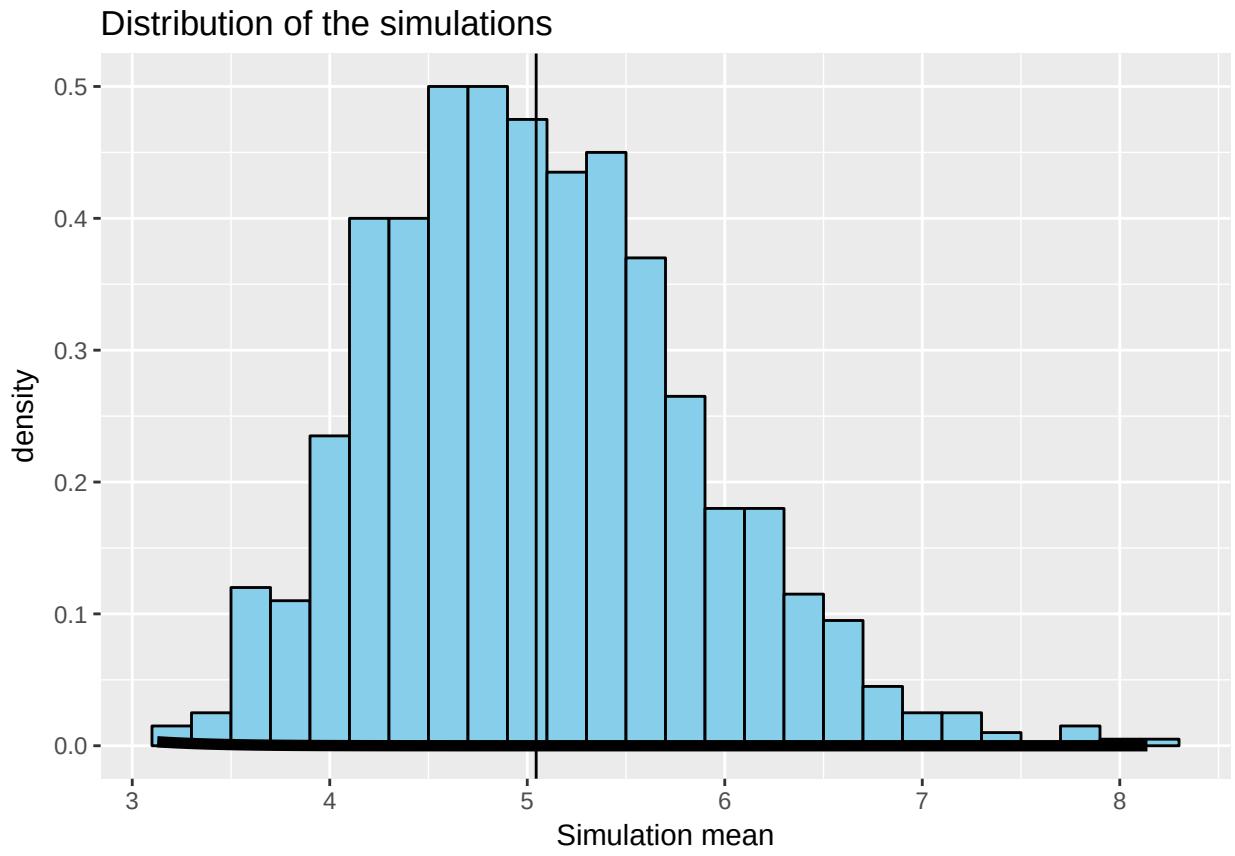
```
## [1] 0.015
```

Answer of demand 3

Show that the distribution is approximately normal:

```
library(ggplot2)

ggplot(data.frame(y=df_sim), aes(x = y)) +
    geom_histogram(aes(y=..density..),
                   binwidth = 0.2,
                   fill='skyblue',
                   color='black') +
    geom_vline(xintercept = sm) +
    stat_function(fun = dnorm,
                  arg = list(mean = sm,
                             sd = sv),
                  size=2) +
    labs(title= "Distribution of the simulations",
         x = "Simulation mean")
```



Conclusions

In this project, we can observe that the sampling exponential distribution is approximately normal when the exponent is equal to 40, and lambda is equal to 0.2.