Part 1: Simulation Exercise

Tomasz Dworowy

```
knitr::opts_chunk$set(warning=FALSE, message=FALSE)
```

```
if (!require("ggplot2"))
  install.packages("ggplot2")
library(ggplot2)
```

Overview

Goal of this rapport is to investigate the exponential distribution in R and compare it with the Central Limit Theorem. The mean of exponential distribution and the standard deviation are both equal to 1/lambda. Lambda is set to 0.2 for all of the simulations. Analysis is done for diminution of averages of 40 exponentials for thousand simulations.

Simulations

Set variables.

```
set.seed(999)
lambda = 0.2
exponentials = 40
n = 1000
```

Run Simulations.

```
means = NULL
for (i in 1 : n) means = c(means, mean(rexp(exponentials, lambda)))
```

Sample Mean versus Theoretical Mean

Sample Mean

Calculating the mean from the simulations with give the sample mean.

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

```
## [1] 5.029028
```

Theoretical Mean

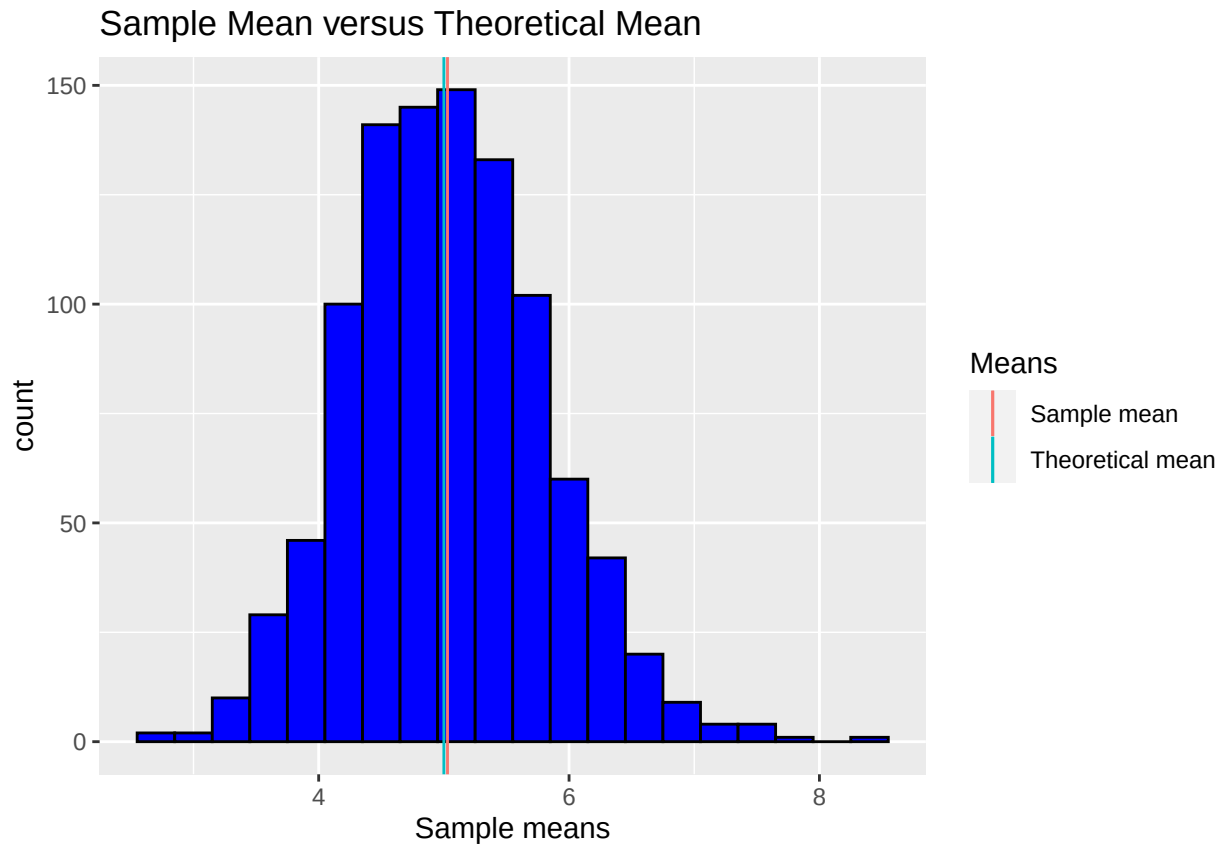
The theoretical mean of an exponential distribution is lambda^-1.

```
theoretical_mean <- lambda^-1
theoretical_mean
```

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

The difference between the sample mean **(5.0290278)** and theoretical mean **(5)** is **0.0290278**

```
ggplot(as.data.frame(means), aes(x=means)) +
     geom_histogram(binwidth = .3, color = "black", fill = "blue") +
     geom_vline(aes(xintercept = sample_mean, color = "green")) +
     geom_vline(aes(xintercept = theoretical_mean, color = "red")) +
     labs(title = "Sample Mean versus Theoretical Mean", x = "Sample means") +
     scale_color_discrete(name = "Means", labels = c("Sample mean", "Theoretical mean"))
```



Sample Variance versus Theoretical Variance

Sample Variance

Calculating the variance from the simulations with give the sample mean.

```
sample_var <- var(means)
sample_var
```

```
## [1] 0.605616
```

Theoretical Variance

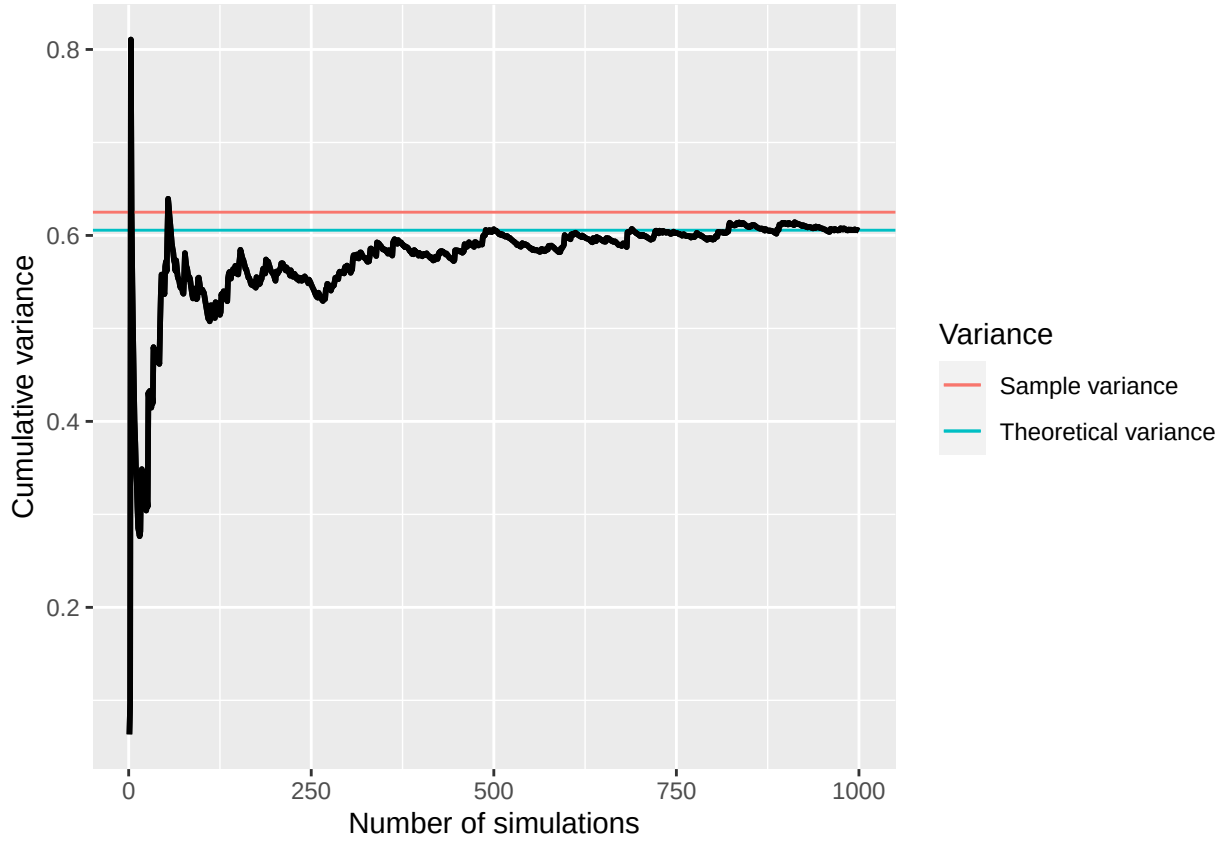
The theoretical variance of an exponential distribution (lambda * sqrt(n))^-2.

```
theoretical_var <- (lambda * sqrt(exponentials))^-2
theoretical_var
```

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

The difference between the sample variance **(0.605616)** and theoretical variance **(0.625)** is **-0.019384**

```
cumvar <- cumsum((means - sample_mean)^2)/(seq_along(means - 1))
ggplot(data.frame(x = 1:n, y = cumvar), aes(x = x, y = y)) +
 geom_hline(aes(yintercept = sample_var, colour = 'red')) + geom_line(size = 1) +
 geom_hline(aes(yintercept = theoretical_var, colour = 'green')) + geom_line(size = 1) +
 labs(x = "Number of simulations", y = "Cumulative variance") +
scale_color_discrete(name = "Variance", labels = c("Sample variance", "Theoretical variance"))
```

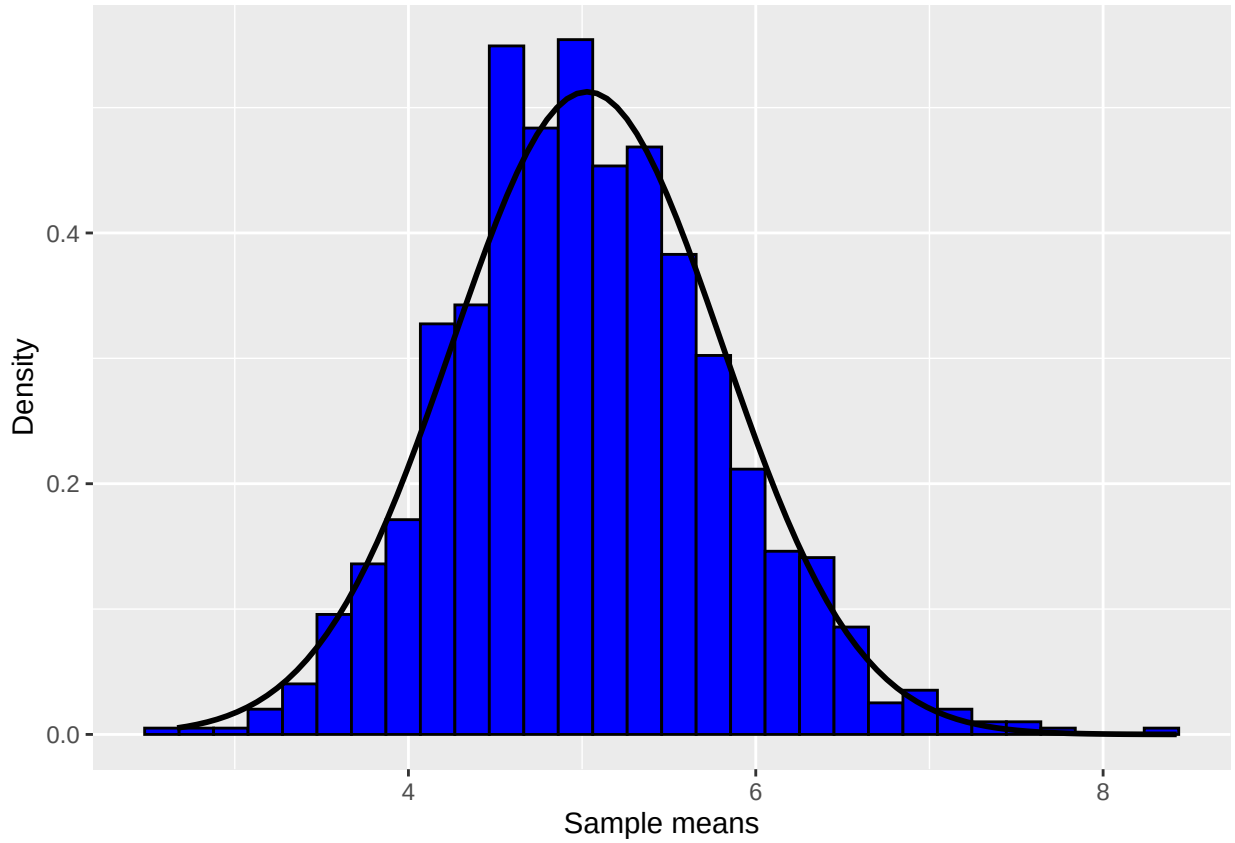


Distribution

Show that the distribution is approximately normal.

```
ggplot(as.data.frame(means), aes(x=means)) +
    geom_histogram(aes(y = ..density..),colour="black",fill="blue")+
```

```
    stat_function(fun=dnorm, args=list( mean=sample_mean, sd=sqrt(sample_var)),geom="line",color = "bla
    scale_x_continuous("Sample means")+
    ylab("Density")
```



As we can see distribution follows raffly the normal bell shape distribution.