Simulation of exponentional distribution

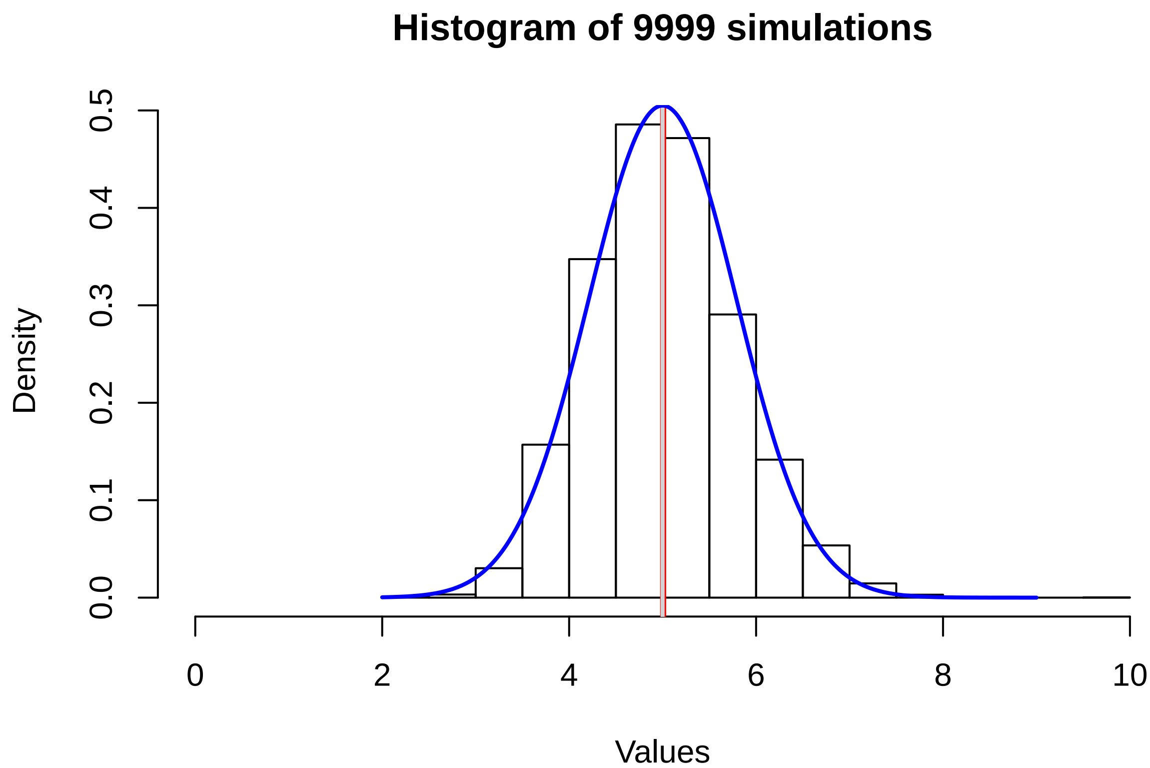
Introduction

In this report we are going to simulate an exponentional distribution with lambda=0.2 and investigate its properties such as mean and variance.

Simulation method

The rate parameter for the exponential distribution $\lambda = 0.2$ and sample size = 40. The sampling were performed 9999 times and the density histogram was produced.

```r
library(knitr)
set.seed(1234)
lambda <- 0.2
n <- 40
trail <- 9999
    mean_trail <- c()
    sd_trail <- c()
    for(i in 1:trail){
    values <- rexp(n, lambda)
    mean <- mean(values)
    sd <- sd(values)
    mean_trail <- c(mean_trail, mean)
    sd_trail <- c(sd_trail, sd)
    }
# histogram plot
hist(mean_trail, freq= FALSE, xlim=c(0,10),
     main=paste("Histogram of",trail,"simulations"), xlab = "Values")
abline(v=mean(mean_trail), col= "red",lwd=3)
abline(v=1/lambda, col="lightgrey", lwd = 2)
lines(seq(2,9,by=0.02), dnorm(seq(2,9,by=0.02),
                              mean=5,sd=0.79),col="blue",lwd=2)
```




```
print(paste("Mean value of the simulation is",
            round(mean(mean_trail),3),", as shown in the graph(red line)"))
```

```
## [1] "Mean value of the simulation is 5.006 , as shown in the graph(red line)"
```

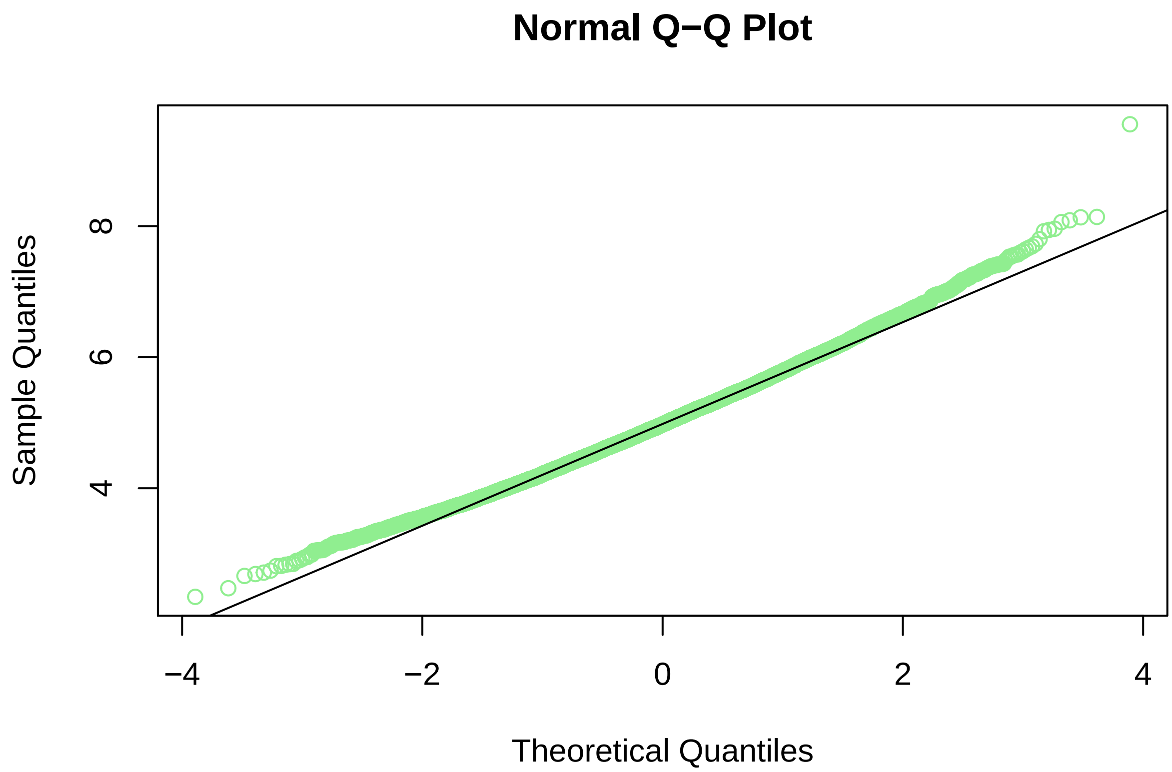
```
print(paste("The standard deviation of the simulation is",
            round(sd(mean_trail),3)))
```

```
## [1] "The standard deviation of the simulation is 0.784"
```

The theoretical center for the simulated exponential distribution is $1/\lambda$, which is 5 in this case. Comparing with the value we obtained 5.01, they are very close. Similarly the theoretical standard deviation is $1/\lambda$, while the simulated one, according to the central limit theorem, should be $1/\lambda$ devided by $\sqrt{n}$, in this case is 0.791, which is also very close to the above value 0.784.

A QQ plot is shown below, as you can see the simulation deviates slightly from the line, indicating the simulation is approximately normal.

```
qqnorm(mean_trail, col="lightgreen")
qqline(mean_trail)
```



The coverage for $\bar{X} \pm 1.96s/\sqrt{n}$ is calculated as the percentage of the times that the true mean value 5 fell in between each confidence interval.

```
upper <- mean_trail + sd_trail*1.96/sqrt(n)
lower <- mean_trail - sd_trail*1.96/sqrt(n)
percent <- sum(lower < 5 & 5 < upper)/10000
print(paste("The coverage for the confidence interval is",percent*100,"%" ))
```

```
## [1] "The coverage for the confidence interval is 92.51 %"
```