Simulation Exercise: Exponential Distributions

Kenneth Fajardo

4/29/2021

Overview

This project investigates the variations of the exponential distribution as the population size approaches a large number 1000. The probability distribution is given by

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

where λ is set to 0.2.

Initialization

The data consists of the means of randomly generated probabilities using the *regexp(n, λ)* function, for a population size of 1000. We used a *for* loop for the calculation. λ is set to 0.2. For consistency, we also set the random seed to 1.

```
set.seed(1)
data <- c()      # create blank vector
n <- 1000        # population size
lambda <- 0.2    # assign lambda

for(i in seq_along(1:1000)){
    # calculate and store the mean of the values randomly generated
    # over the exponential distribution
    data <- rbind(data, mean(rexp(n, lambda)))
}
```

1. Show the sample mean and compare it to the theoretical mean of the distribution.

The theoretical mean μ in an exponential distribution is given by

$$\mu = \frac{1}{\lambda}$$

Substituting the value of λ, we get $\mu = 5$

```
1/lambda
```

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

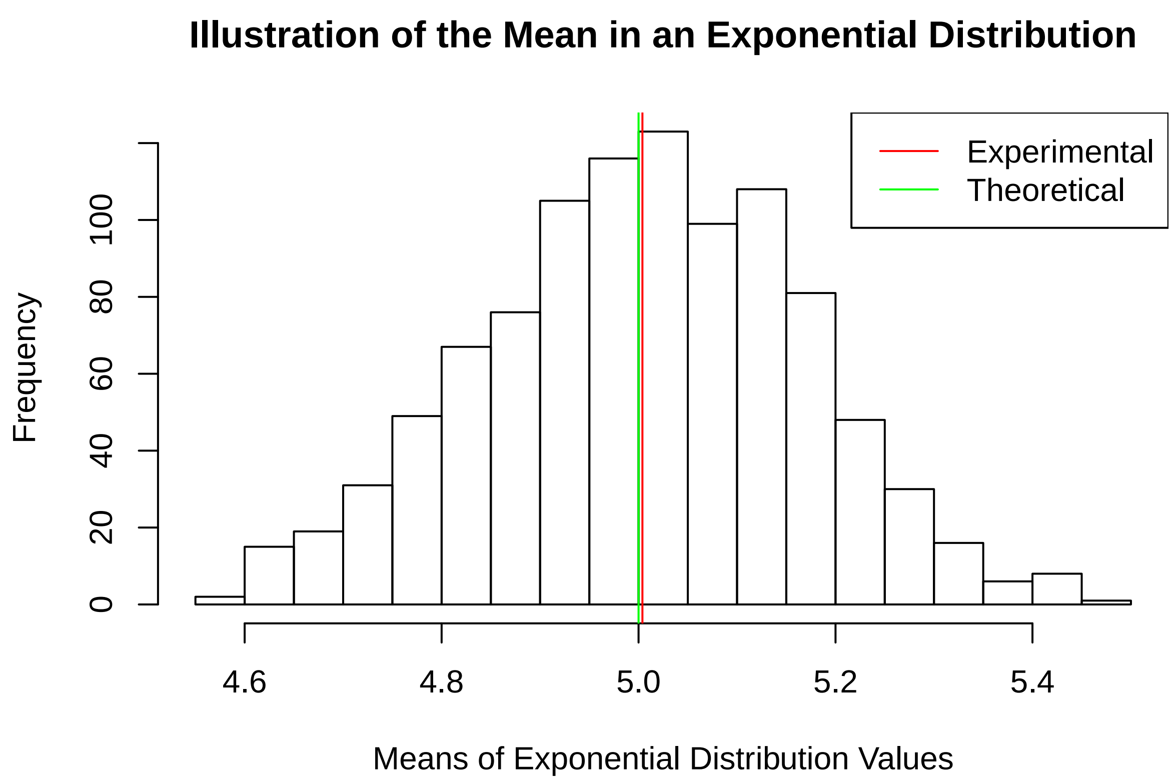
Now, for calculating the sample mean, we only need the *mean* function in R.

```
mean(data)
```

```
## [1] 5.00393
```

We see that **the sample mean is approximately equal to the theoretical mean** ($\mu \approx 5$). To further illustrate, see the figure below.

```
hist(data, breaks=20, main="Illustration of the Mean in an Exponential Distribution",
          xlab="Means of Exponential Distribution Values", ylab="Frequency")
abline(v=mean(data), col="red")
abline(v=1/lambda, col="green")
legend("topright", legend=c("Experimental","Theoretical"), col=c("red", "green"), lty=1)
```



2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution. The theoretical variance σ^2 in an exponential distribution is given by

$$\sigma^2 = \frac{1}{\lambda^2 n}$$

Substituting the value of λ, we get $\sigma^2 = 25$

```
1/(lambda^2 *n)
```

```
## [1] 0.025
```

R has a built-in *var* function for calculating variance. Using it on our data, we get

```
var(data)
```

```
##              [,1]
## [1,] 0.02667744
```

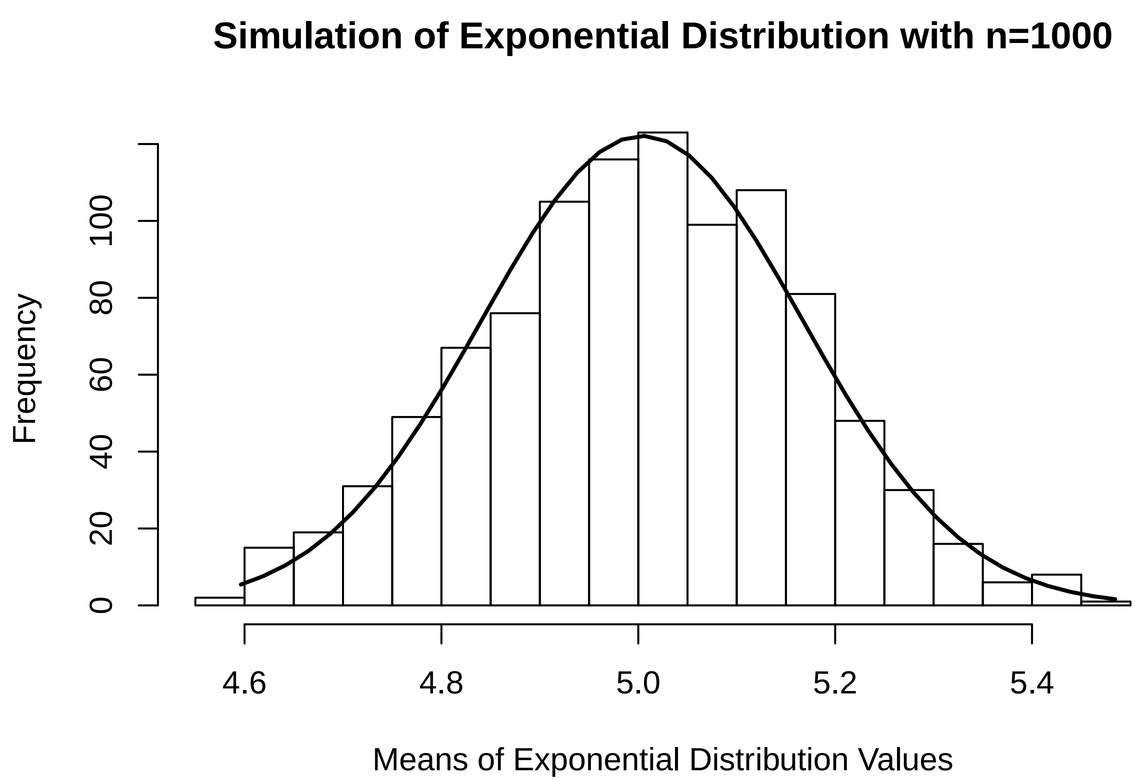
We can see that the **theoretical variance is approximately equal to the sample variance** ($\sigma^2 \approx 0.025$).

3. Show that the distribution is approximately normal.

One way to show normality is through histograms. Using the function *hist*, we can see the distribution of the means of exponential values. We, then, overlay the histogram with a bell curve, representing normal distribution.

```
h <- hist(data, breaks=20, main="Simulation of Exponential Distribution with n=1000",
          xlab="Means of Exponential Distribution Values", ylab="Frequency")
xfit <- seq(min(data), max(data), length = 40)
yfit <- dnorm(xfit, mean = mean(data), sd = sd(data))
yfit <- yfit * diff(h$mids[1:2]) * length(data)

lines(xfit, yfit, col = "black", lwd = 2)
```



We can observe that the **frequencies of the means follow a normal distribution**. This is true according to the Central Limit Theorem when n approaches large numbers.