Exponential distribution in R and Central Limit Theorem. Comparission

Miguel Angel Vallejo Gato

25 de octubre de 2015

Contents

1. Overview

In this project we are going to investigate the exponential distribution in R and will compare it with the Central Limit Theorem.

We will show: 1. The sample mean and compare it to the theoretical mean of the distribution. 2. How variable the sample is (via variance) and compare it to the theoretical variance of the distribution. 3. That the distribution is approximately normal.

2. Simulations

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.
- The standard deviation is also 1/lambda.

For the simulation, we are going to set the next values:

- lambda = 0.2
- Number of exponentials for the distribution of averages = 40.
- Number of simulations = 1000.

```
lambda = 0.2
exponentials = 40
simulations = 1000
```

We are going to set a seed to make the experiment reproducible.

```
set.seed(1)
```

Finally, We are going to create a vector (means) with the means. In order to do that, we will repeat 1000 times the next steps:

1. We wil generate a ramdon exponential distribution rexp(exponentials,lambda)
2. We will calculate the mean
3. We will add the value to the means vector.

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

3. Sample Mean versus Theoretical Mean

The theoretical mean of the exponential distribution is 1/lambda

```
1/lambda
```

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

The mean of the means vector values is:

```
mean(means)
```

```
## [1] 4.990025
```

As we can see, the sample mean is similar to the theoretical mean of the distribution (1/lambda=5)

4. Sample Variance versus Theoretical Variance

The standard deviation of exponential distribution is 1/lambda.

The variance of an exponential distribution is ((1/lambda)^2)/exponentials

```
tVar <- round(((1/lambda)^2)/exponentials,3)
```

Now we are going to calculate the variance of the means of the samples. We will use the R "var" function.

```
sVar <- round(var(means),3)
```

Let's show the values:

```
cbind(sVar,tVar)
```

```
##       sVar  tVar
## [1,] 0.611 0.625
```

As we can see, the variance of the sample means is very similar to the Theorical Variance.

5. Distribution

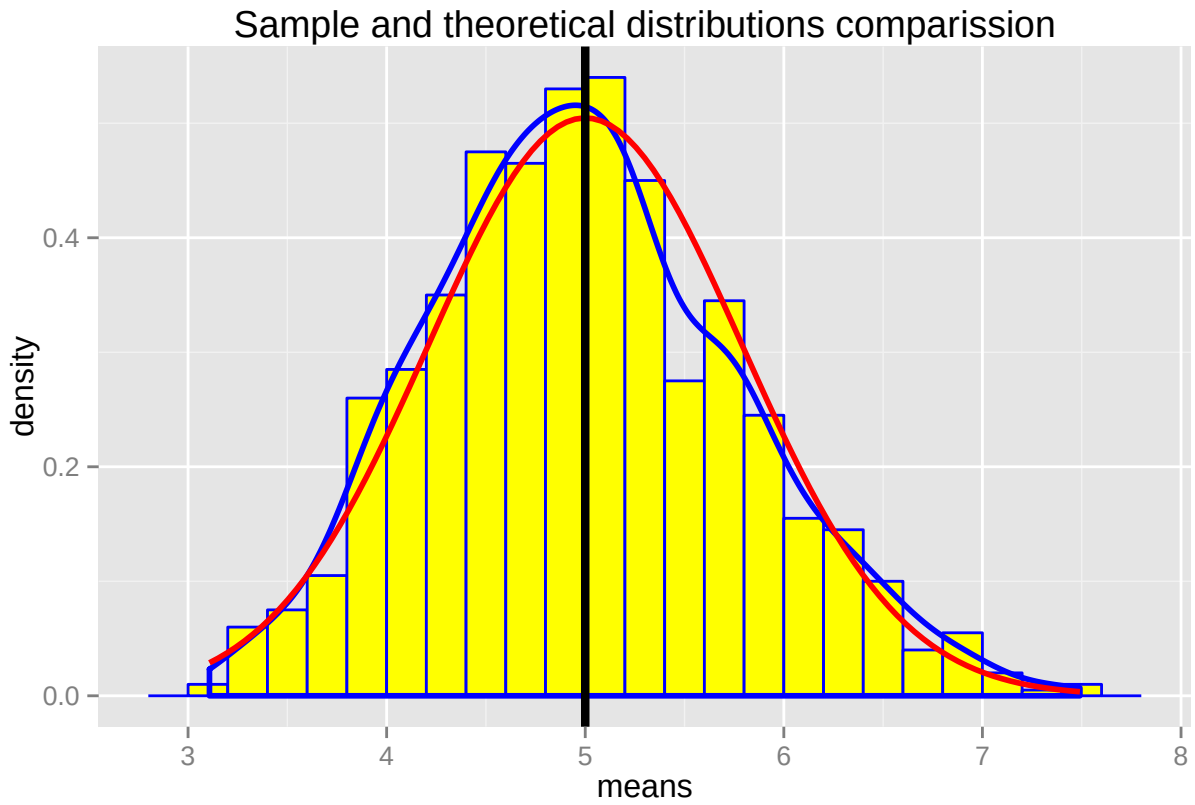
We are going to draw a new plot with:

1. The means distribution.
2. A curve (blue) showing the means density.
3. A curve (red) showing the density of a normal distribution.
4. A line showing the mean.

We will use the "ggplot2"" R library to make the plot.

```
library(ggplot2)

ggplot(data = as.data.frame(means), aes(x = means)) +
    ggtitle("Sample and theoretical distributions comparission") +
    geom_histogram(binwidth=0.2, aes(y=..density..),col="blue",fill="yellow") +
    geom_density(colour="blue", size=1) +
    stat_function(fun = dnorm, arg = list(mean = 1/lambda , sd = sqrt(tVar)),
                  colour = "red", size=1) +
    geom_vline(xintercept = 5, size=1.5, colour="black")
```



We can observe that the means population is similar to a normal distribution.