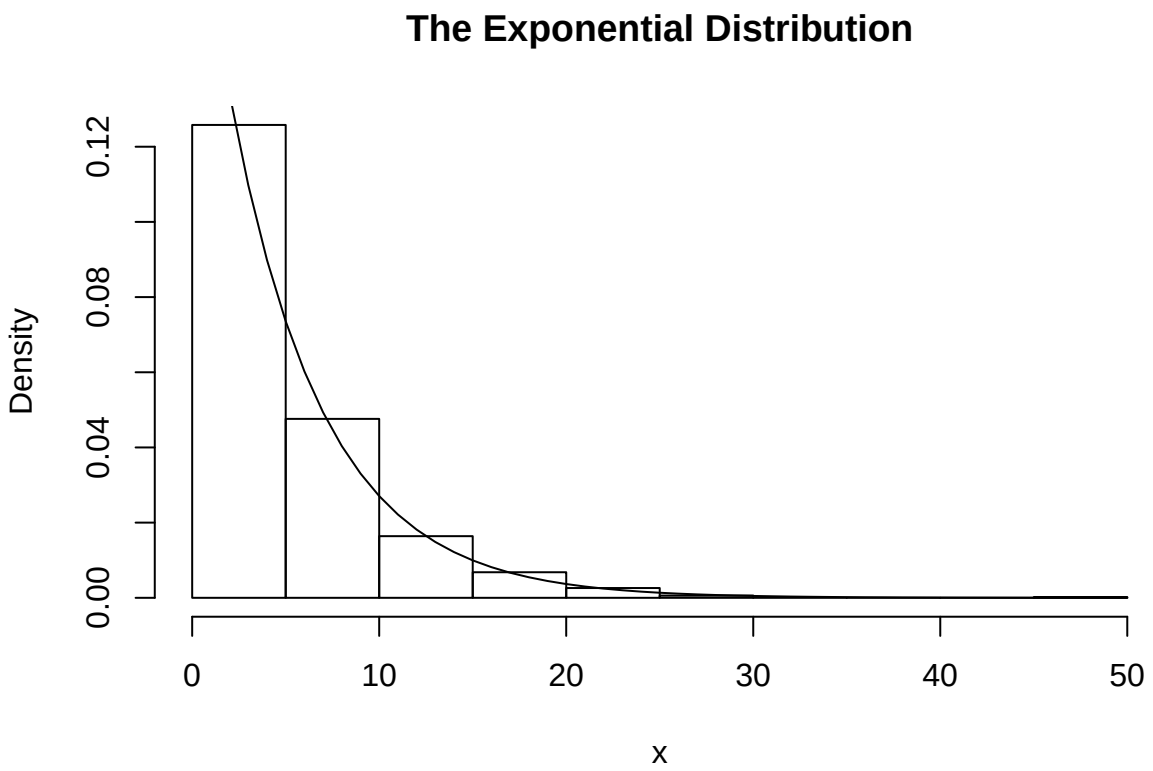
The Exponential Distribution in the Central Limit Theorem

Sebastien Viaene

18 Jan 2016

In this project we investigate the exponential distribution in R and compare it with the Central Limit Theorem. The exponential distribution is defined by the lambda parameter. Theoretically, the mean of exponential distribution is 1/lambda and the standard deviation is also 1/lambda. We set lambda = 0.2 for all of the simulations. This is what our exponential distribution looks like:

```
lambda <- 0.2
# Show an example of the exponential distribution
hist(rexp(1000, lambda),freq=FALSE, xlab='x', main="The Exponential Distribution")
x <- seq(0,50)
lines(x,lambda*exp(-lambda*x))
```



Here, the histogram shows the distribution of 1000 random numbers drawn form the exponential distribution, and the solid line shows the theoretical curve for the distribution.

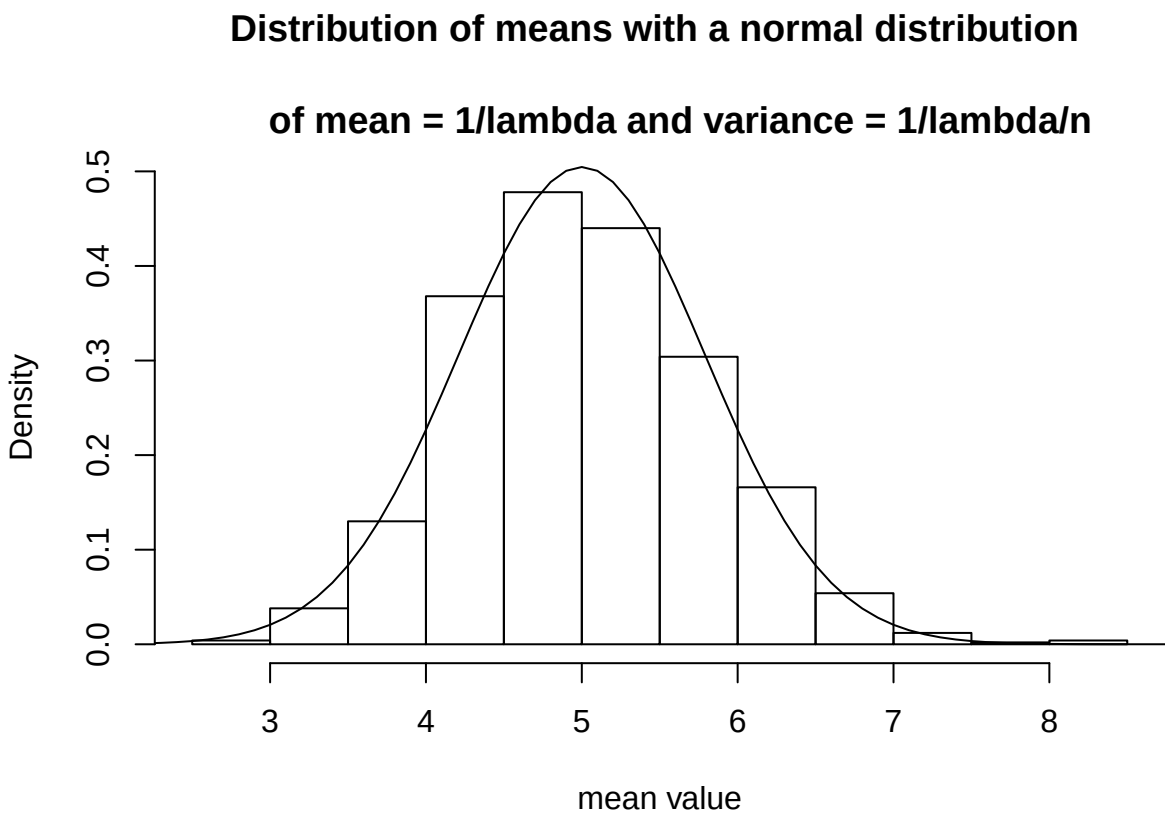
Now lets simulate 40 values from the exponential distribution and take their mean. If we do that 1000 times, we can plot the distribution of the means.

```
# Sample size
n <- 40
# Number of simulations
nsim <- 1000
```

```r
# Matrix of values drawn from exponential distribution
TheMatrix <- matrix(rexp(nsim * n, lambda), nsim)
# Apply mean over rows
meansample <- apply(TheMatrix,1 , mean)

# Plot the distribution of the means
hist(meansample,freq=FALSE, xlab='mean value',
     main="Distribution of means with a normal distribution \n
     of mean = 1/lambda and variance = 1/lambda/n ",ylim = c(0.0,0.5))

# Now add a Gaussian to this
xnorm <- seq(-10,10,0.1)
lines(xnorm, dnorm(xnorm, mean=1/lambda, sd=1/lambda/sqrt(40)))
```



On this figure, we have overplotted a normal distribution with a theoretical mean of 1/lambda, which should be the same as the mean of the exponential. The standard deviation of this normal distribution is set to be 1/lambda/sqrt(n) where n = 40, the sample size for each mean. This value gives the standard error on the exponential distribution. From a first look at the histogram, it appears that it corresponds well to the normal distribution. Let's have a look at the mean and the variance of the histogram.

```r
sampleMean <- mean(meansample)
sampleVar <- var(meansample)
```

We find that the mean of the distribution of means is 5.0297802, this is very close to the theoretical value of 1/lamdba = 1/0.2 = 5.0. For the variance of the sample of means, we find 0.6376448. This should be compared to the theoretical value of 1/lambda^2/n, or 0.625. Again, the values lie close together.

Our analysis suggests that the Central Limit Theorem applied to the exponential distribution.