Exponential Distribution in R vs Central Limit Theorem

Stefaan Delanghe

30 June 2016

In this project you will investigate the exponential distribution in R and compare it with the Central Limit Theorem.We will investigate the distribution of averages of 40 exponentials. A thousand simulations will be performed.

```
## Warning: package 'ggplot2' was built under R version 3.2.4
```

Simulations

Declaration of the simulation requirements.

```
nrsim <- 1000
lambda <- 0.2
nrexp <- 40
distmeans <- NULL
for (i in 1 : nrsim) distmeans <- c(distmeans, mean(rexp(nrexp, lambda)))
```

Sample Mean versus Theoretical Mean

Investigation how the sample mean is related to the theoretical mean

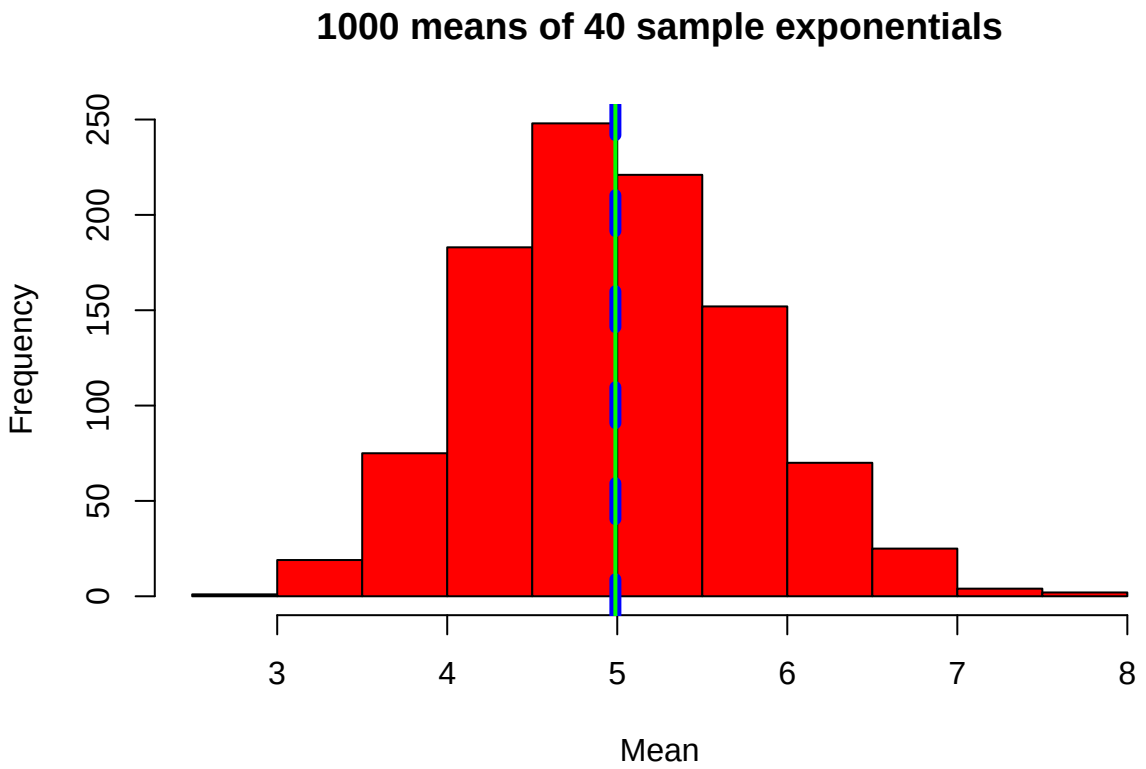
```
sample.mean <- round(mean(distmeans),3);
sample.mean
```

```
## [1] 4.989
```

```
theoretical.mean <- round(1/lambda,3);
theoretical.mean
```

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

```
hist(distmeans, main="1000 means of 40 sample exponentials", col="red", xlab = "Mean")
abline(v = sample.mean, col = "blue", lwd = 6, lty=2)
abline(v = sample.mean, col = "green", lwd = 2)
```



The histogram shows the sample (blue) and theoretical (green) mean. The sample mean is 4.989 while the theoretical mean is 5. This means that the sample mean is very close to the theoretical mean.

Sample Variance versus Theoretical Variance

Investigation how the sample variance is compared to the theoreticam variance.

```
theoretical.variance <- ((1/lambda)^2) / nrexp
theoretical.variance
```

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

```
mean.variance <- var(distmeans)
mean.variance
```

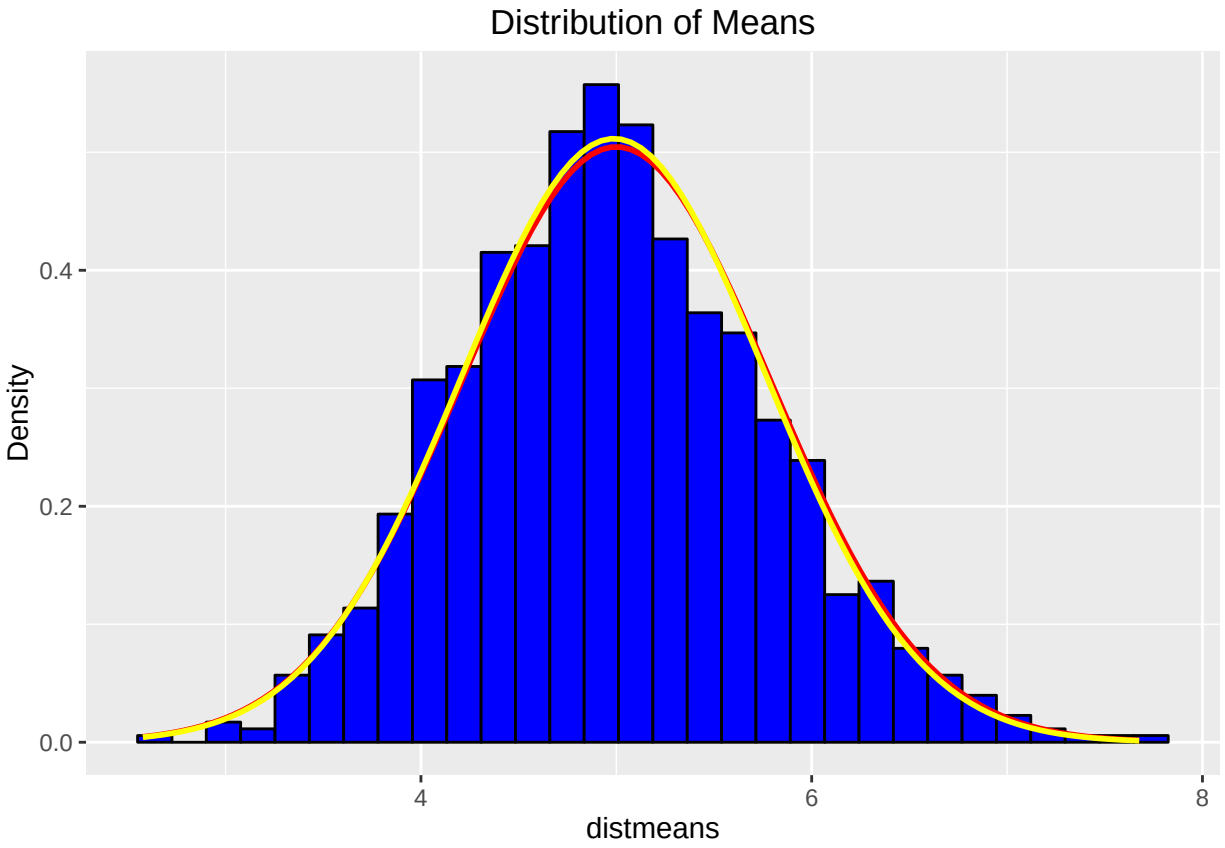
```
## [1] 0.6081224
```

The theoretical variance of the distribution is 0.625 while the the sample vaiance of the distribution is 0.6081224. The difference between the theoretical and sample variance is very small.

Distribution

Investigate if the distribution is approximately normal.

```r
p.data <- data.frame(distmeans)
p <- ggplot(p.data, aes(x = distmeans))
p <- p + geom_histogram(aes(y=..density..), colour="black", fill="blue", bins = 30)
p <- p + labs(title="Distribution of Means", y="Density", xlab="Means")
p <- p + stat_function(fun=dnorm,args=list( mean=1/lambda, sd=sqrt(theoretical.variance)),color = "red"
p <- p + stat_function(fun=dnorm,args=list( mean=mean(distmeans), sd=sqrt(mean.variance)),color = "yell
print(p)
```



The red line is the theoretical variance as the yellow line is the mean variance. The two lines are pretty close aligned with each other. This close alignement indicates that the distribution is approximately normal.