Statistical Analysis

S.McDowell

8/13/2020

Load library's

```
library(ggplot2)
library(tinytex)
```

Overview

This project you will investigate the exponential distribution in R and compare it with the Central Limit Theorem.

Simulations

```
set.seed(12344321)
#Means
mns = NULL
lambda <- 0.2
n <- 40
for (i in 1:1000) mns = c(mns, mean(rexp(n,rate = lambda)))
```

Sample Mean versus Theoretical Mean

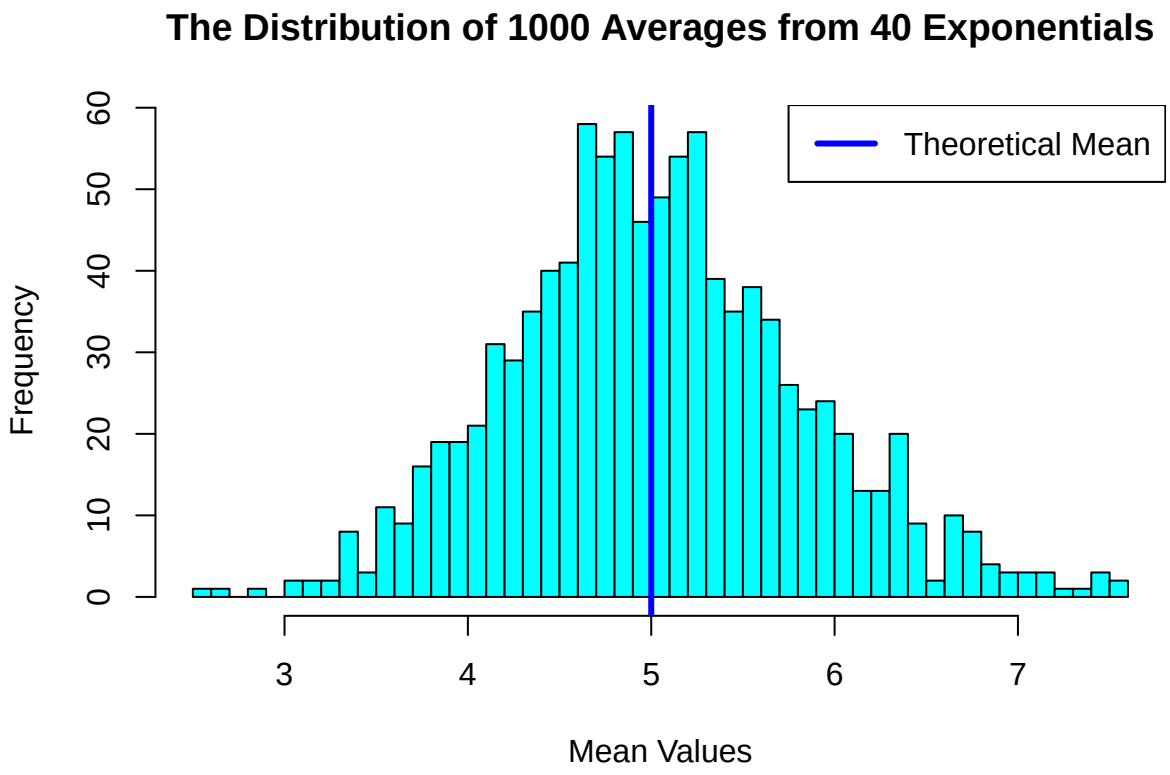
```
#Means
hist(mns,breaks = 40,col = "cyan", main = "The Distribution of 1000 Averages from 40 Exponentials",xlab
sample_mean <- mean(mns)
sample_mean
```

```
## [1] 5.024544
```

```
th_mean<- 1/lambda
th_mean
```

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

```
abline(v = th_mean, col="blue",lwd = 3)
legend("topright",col = "blue", lty = 1, lwd = 3, legend = "Theoretical Mean" )
```



We see that the sample mean and the theoretical mean are very similar.

Sample Variance versus Theoretical Variance

```
#Variances
exp_sd <- (1/lambda)/sqrt(n)
paste("The theoretical standard deviation is",exp_sd, sep = ",")
```

```
## [1] "The theoretical standard deviation is,0.790569415042095"
```

```
exp_var <- (exp_sd)^2
paste("The theoretical variance is",exp_var, sep = ",")
```

```
## [1] "The theoretical variance is,0.625"
```

```
sample_sd <- sd(mns)
paste("The sample standard deviation is",sample_sd, sep = ",")
```

```
## [1] "The sample standard deviation is,0.80026477497704"
```

```
sample_var <- var(mns)
paste("The sample standard variance is",sample_var, sep = ",")
```

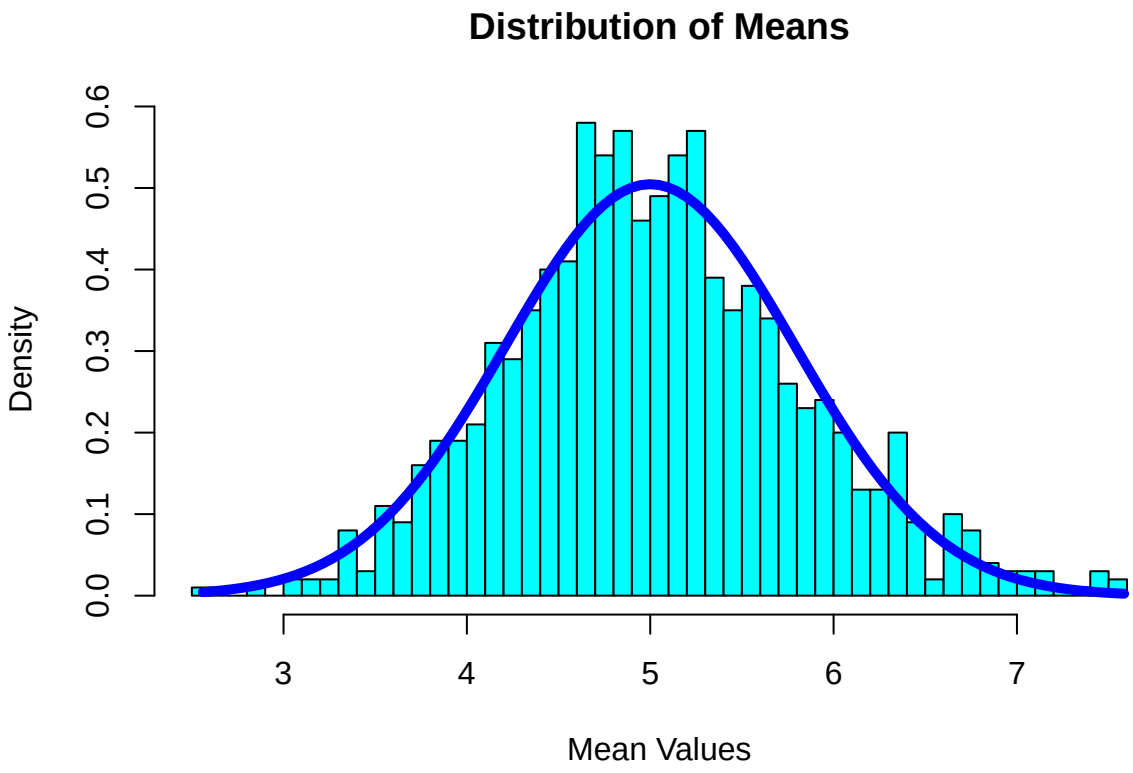
```
## [1] "The sample standard variance is,0.640423710069052"
```

We see that the sample standard deviation and sample variance are close approximations of the theoretical standard deviation and variance.

Distribution

```
fit <- seq(min(mns),max(mns), length=100)
standard_fit <- dnorm(fit,mean =th_mean,sd=exp_sd )

hist(mns,breaks = n,col = "cyan",prob=T, main = "Distribution of Means",xlab = "Mean Values")
lines(fit, standard_fit,pch = 2, col="blue",lwd=5)
```



Here we have the distribution of means overlayed with an normal cure indication that the simulation means are normally distributed.