

StatInf-Project2

Miguel Porro

October 22, 2018

Statistical Inference Project

Overview

The purpose of this analysis is to investigate the exponential distribution and compare the same to the Central Limit Theorem. The parameter (i.e., lambda) will be set to 0.2 for all the simulations. This analysis will compare the distribution of averages of 40 exponentials.

Simulations

Setting seed (3006) and parameter values first, this code will generate a distribution of averages of 40 exponentials for 1000 simulations.

```
set.seed(3006)
lambda <- 0.2
nexp <- 40
average <- NULL
for(i in 1:1000)
  average <- c(average, mean(rexp(nexp, lambda)))
```

Sample Mean versus Theoretical Mean

Theoretical mean of distribution

```
lambda ^ -1
```

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

Sample mean of distribution

```
mean(average)
```

```
## [1] 5.012704
```

Both the theoretical and the sample mean are approximately the same, the difference is not big. ## Sample Variance versus Theoretical Variance

Theoretical variance of distribution

```
(lambda * sqrt(nexp)) ^ -2
```

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

Sample variance of distribution

```
var(average)
```

```
## [1] 0.6504656
```

Again it can be seen that both the theoretical and the sample variance are approximately the same.

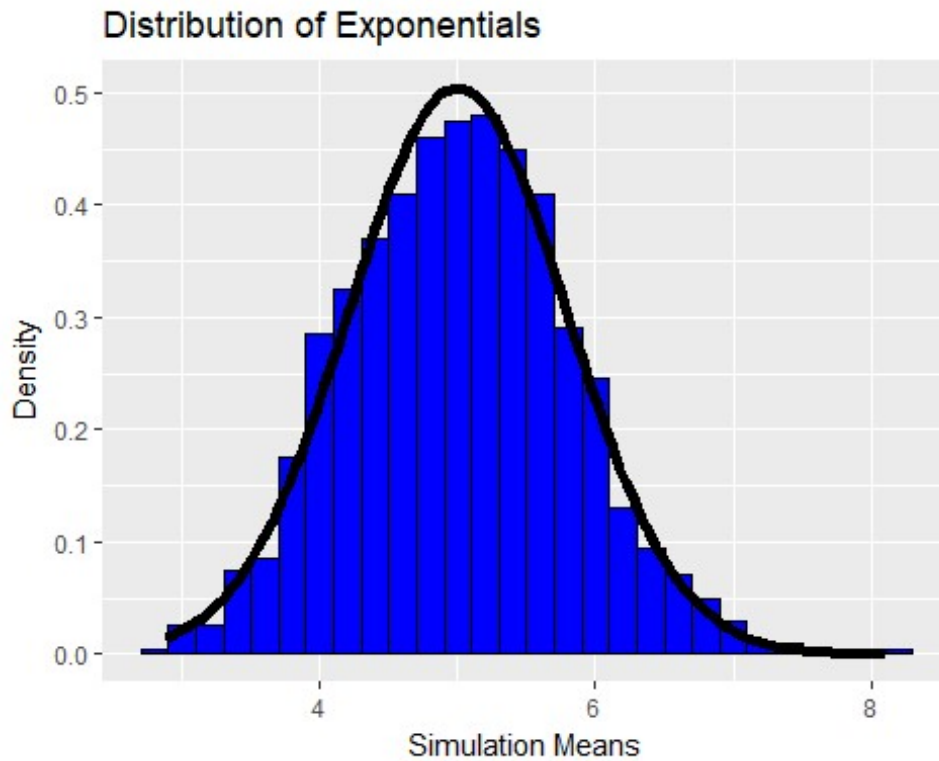
Distribution

Given below is the density histogram of the 1000 simulations. Added over the same plot is the normal distribution with mean (λ^{-1}) and variance of $((\lambda \cdot \sqrt{n_{exp}})^{-2})$ which are the theoretical parameter values for the normal distribution occurring due to the simulations.

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.4.4
```

```
g <- ggplot(data.frame(column = average), aes(x = column))
g <- g + geom_histogram(aes(y = ..density..), binwidth = 0.2, fill =
'blue', color = 'black')
g <- g + stat_function(fun = dnorm, args = list(mean = lambda^-1,
sd=(lambda*sqrt(nexp))^-1), size=2)
g <- g + labs(title = "Distribution of Exponentials", x = "Simulation
Means", y = "Density")
g
```



As from the plot it can be seen that the normal density distribution overlayed onto the density histograms fits perfectly and hence shows that the distribution of averages of a large sample of exponentials is indeed normal.