Comparison of Exponential Distribution of large sample size with Central Limit theorem

B Santhoshini

2023-03-27

Title

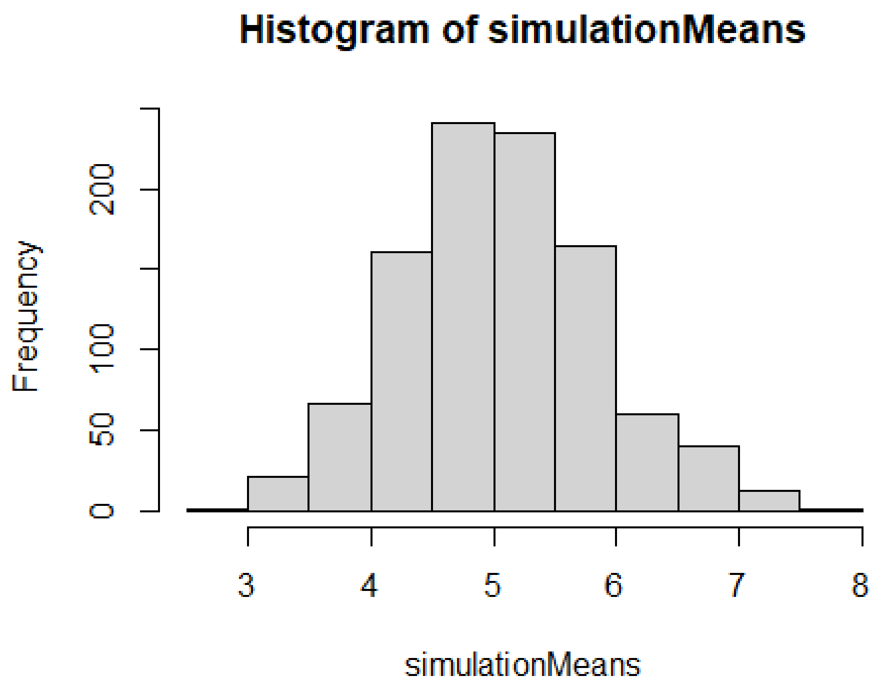
Comparison of Exponential Distribution of large sample size with Central Limit theorem

Overview

In this data analysis we are going to compare the exponential distribution with Central Limit theorem. For this analysis we assume lambda as 0.2 for all the simulations. Here we are going to compare the distribution of averages of 40 exponentials over 1000 simulations #Simulation Variables exponentials = 40, lambda = 0.2

```
ECHO=TRUE
set.seed(1337)
lambda = 0.2
exponentials = 40

simulationMeans = NULL
for (i in 1:1000) {
  simulationMeans = c(simulationMeans,mean(rexp(40,0.2)))
}
hist(simulationMeans)
```



##Embedding plots

##Simulations Above we generated the distribution of 40 averages of random exponentials over 100 simulations.We named a variable simulationMeans and gave it an initial value NULL. Then using a for loop we generated averages of 40 exponentials over 1000 simulation. Then by using a function called hist in R for visualizing data in the form of histogram we made a plot of simulation Means. ##Comparison of Means ## Sample Mean

```
smean <- mean(simulationMeans)
smean

## [1] 5.055995
```

Theoretical Mean = 1/lambda

```
lambda <- 0.2
mean <- lambda^-1
mean

## [1] 5
```

Sample Mean Versus Theoretical Mean

There is only a slight difference between the sample mean and the theoretical mean which corroborates the idea that if the sample size of the distribution is large enough then the exponential distribution or any type of distribution follows normal/Gaussian distribution. The very basic idea of Central Limit Theorem.

```
abs(mean - smean)
```

```
## [1] 0.05599526
```

##Comparison of Variance

##Sample Variance

```
svar<-var(simulationMeans)
svar

## [1] 0.6543703
```

##Theoretical Variance Variance of an exponential distribution is equal to (lambda * sqrt(n))^-2 exponentials(n) = 40 so,

```
var <- (lambda * sqrt(40))^-2
var

## [1] 0.625
```

Sample Variance Versus Theoretical Variance

```
abs(var - svar)

## [1] 0.0293703
```

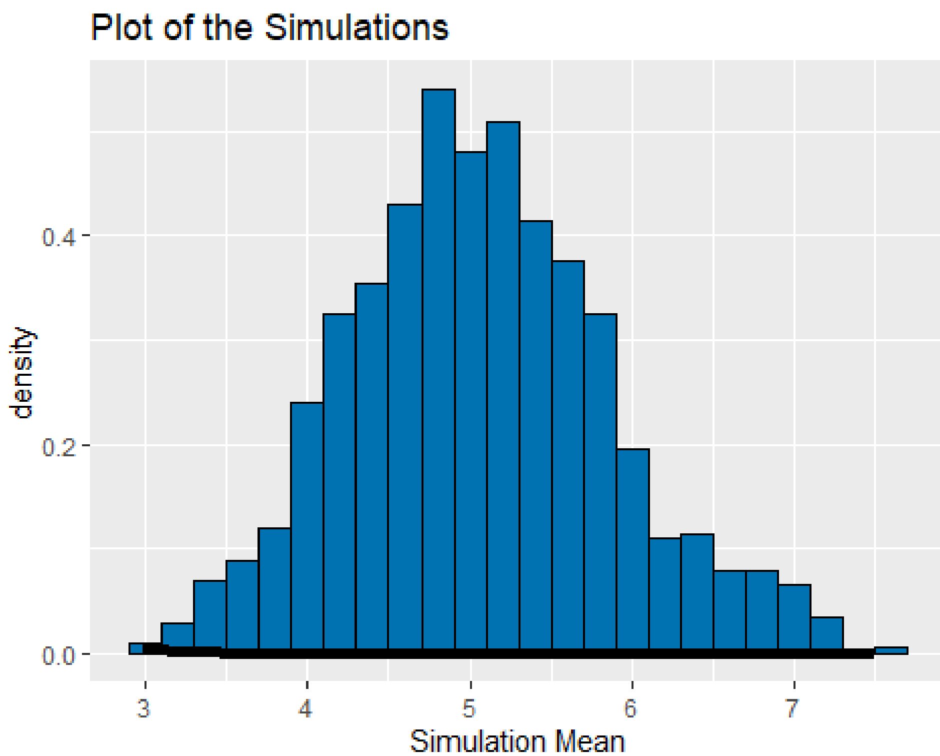
There is only a slight difference between the Sample Variance and the Theoretical Variance ##Distribution

```
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.2.1

ggplot(data.frame(y=simulationMeans), aes(x=y)) +
  geom_histogram(aes(y=..density..), binwidth=0.2, fill="#0072B2",
                 color="black") +
  stat_function(fun=dnorm, arg=list(mean=lambda^-1,
                                    sd=(lambda*sqrt(40))^-1),
                size=2) +
  labs(title="Plot of the Simulations", x="Simulation Mean")

## Warning: Ignoring unknown parameters: arg
```



Here is a density histogram of 1000 simulations with an over lay of normal distribution with mean lambda^-1 and standard deviation of lambda * sqrt(40).