On the Topic of the Exponential Distribution with respect to the Central Limit Theorem in R

Evan Falcone

20 January, 2018

Overview

This data analysis aims to investigate the Exponential distribution with respect to the Central Limit Theorem under certain constraints. We will use lambda = 0.2 for all simulations. We compare the distribution of averages of 40 exponentials over 1000 simulations.

Simulations

Start by initializing the common variables (lambda, exponentials) and setting a seed:

```
set.seed(1337)
lambda <- 0.2
exponentials <- 40
```

Run simulations with the initialized variables:

```
simMean <- NULL
for(i in 1:1000) {simMean <- c(simMean, mean(rexp(exponentials, lambda)))}
```

Comparison: Sample Mean to Theoretical Mean

Sample Mean

We compute the sample mean by taking the mean of the simulations:

```
mean(simMean)
```

```
## [1] 5.055995
```

Theoretical Mean

The theoretical mean of an exponential distribution is given by lammbda ^ -1:

```
lambda^-1
```

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

Comparison

The difference between the simulations' sample mean and the theoretical mean of the exponential distribution is small (order of magnitude = 10^-2):

```
abs(mean(simMean)-(lambda^-1))
```

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

Comparison: Sample Variance to Theoretical Variance

Sample Variance

We compute the sample variance by finding the variance of the simulations:

```
var(simMean)
```

```
## [1] 0.6543703
```

Theoretical Variance

The theoretical variance of an exponential distribution is given by (lammbda * sqrt(n))^-2:

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

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

Comparison

The difference between the simulations' sample variance and the theoretical variance of the exponential distribution is small (order of magnitude = 10^-2):

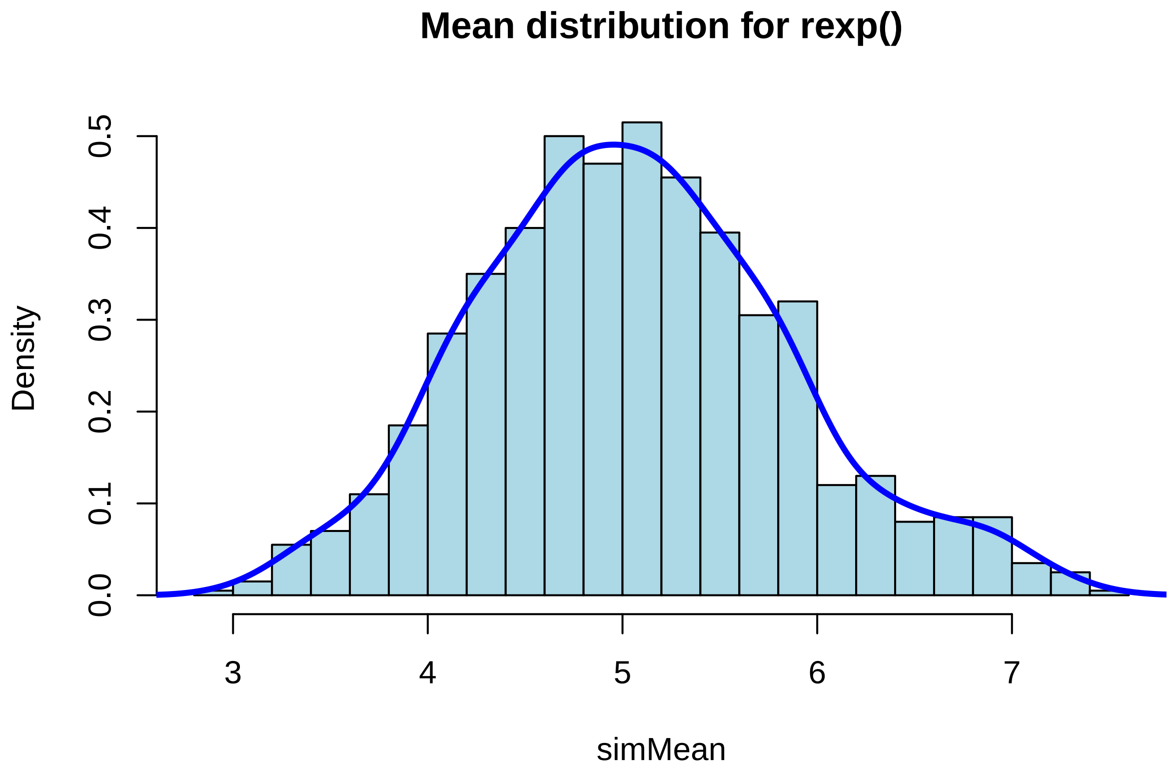
```
abs(var(simMean)-(lambda * sqrt(exponentials))^-2)
```

```
## [1] 0.0293703
```

Distribution

The following is a density histogram plot of the 1000 simulations made using our initialized parameters lambda = 0.2 and seed 1337. We overlay a normal distribution that has a mean == mean(simMean) (i.e. lamda^-1) and variance == var(simMean) (i.e. (lambda * sqrt(exponentials))^-2), the theoretical normal distribution for the simulations we made. Graphically, the two distributions appear to bear a resemblance:

```
hist(simMean, prob=TRUE, col="lightblue", main="Mean distribution for rexp()", breaks=20)
lines(density(simMean), lwd=3, col="blue")
```


Mean distribution for rexp()
Density
0.0
0.1
0.2
0.3
0.4
0.5
3
4
5
6
7
simMean