

Investigate the exponential distribution in R and compare it with the CLT

Philip Wong

29/05/2022

Overview

In this project we will investigate the exponential distribution in R and compare it with the Central Limit Theorem. The exponential distribution can be simulated in R with `rexp(n, lambda)` where `lambda` is the rate parameter. The mean of exponential distribution is $1/\lambda$ and the standard deviation is also $1/\lambda$. Set `lambda = 0.2` for all of the simulations. We will investigate the distribution of averages of 40 exponentials. Note that it will require 1000 simulations.

Simulations

Set the simulation variables `lambda`, `exponentials`, and `seed`.

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

Run Simulations with variables

```
simulation_Means = NULL
for (i in 1 : 1000) simulation_Means = c(simulation_Means, mean(rexp(exponentials, lambda)))
```

Sample Mean versus Theoretical Mean

```
mean(simulation_Means)
```

Sample Mean

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

```
lambda^-1
```

Theoretical Mean

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

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

Comparison

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

Sample Variance versus Theoretical Variance

```
var(simulation_Means)
```

Sample Variance

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

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

Theoretical Variance

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

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

Comparison

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

Distribution

```
library(ggplot2)
ggplot(data.frame(y=simulation_Means), aes(x=y)) +
  geom_histogram(aes(y=..density..), binwidth=0.2, fill="dark green",
                 color="black") +
  stat_function(fun=dnorm, arg=list(mean=lambda^-1,
                                   sd=(lambda*sqrt(exponentials))^-1),
               size=2) +
  labs(
    title = "Plot of the Simulations",
    x = "Simulation Mean")
```

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

Plot of the Simulations

