Investigate the exponential distribution in R and compare it with the Central Limit Theorem

Niveditha Sanapala

October 16 2023

# Overview

The purpos e of this data analys is is to inv estigate the exponential dis tribution and c ompare it to the Central Limit Theorem. Forthis analysis, thelambdawill be setto0.2 for all of thesimulations. This investigation will compare the distribution of averages of 40 exponentials over 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

simMeans = NULL

**for** (i **in** 1 : 1000) simMeans = c(simMeans, mean(rexp(exponentials, lambda)))

# Sample Mean versus Theoretical Mean

## Sample Mean

Calculating the mean from the simulations with give the sample mean.

mean(simMeans)

## [1] 5.055995

## Theoretical Mean

The theoretical mean of an exponential distribution is lambda^-1.

lambda^-1

## [1] 5

## Comparison

There is only a slight difference between the simulations sample mean and the exponential distribution theoretical mean.

abs(mean(simMeans)-lambda^-1)

## [1] 0.05599526

# Sample Variance versus Theoretical Variance

## Sample Variance

Calculating the variance from the simulation means with give the sample variance.

var(simMeans)

## [1] 0.6543703

## Theoretical Variance

The theoretical variance of an exponential distribution is (lambda \* sqrt(n))^-2.

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

## [1] 0.625

## Comparison

There is only a slight difference between the simulations sample variance and the exponential distribution theoretical variance.

abs(var(simMeans)-(lambda \* sqrt(exponentials))^-2)

## [1] 0.0293703

# Distribution

This is a d ens ity his togram of the 1000 simulations . There is an ov erlay with a normal d istribution that has a mean of lambda^ -1 and stand ard d ev iation of (lambda\*sqrt(n))^-1, the theoretic al normal d istribution for the simulations.

**library**(ggplot2) ggplot(data.frame(y=simMeans), 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(exponentials))^-1),

size=2) +

labs(title="Plot of the Simulations", x="Simulation Mean")

