Statistical\_Inference\_project

Maiia Vasileva

2024-07-15

## Instructions

1. Show the sample mean and compare it to the theoretical mean of the distribution.
2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.
3. Show that the distribution is approximately normal.

## Loading Libraries

library("data.table")  
library("ggplot2")

## Task

# set seed for reproducability  
set.seed(31)  
  
# set lambda to 0.2  
lambda <- 0.2  
  
# 40 samples  
n <- 40  
  
# 1000 simulations  
simulations <- 1000  
  
# simulate  
simulated\_exponentials <- replicate(simulations, rexp(n, lambda))  
  
# calculate mean of exponentials  
means\_exponentials <- apply(simulated\_exponentials, 2, mean)

## Question 1

Show where the distribution is centered at and compare it to the theoretical center of the distribution.

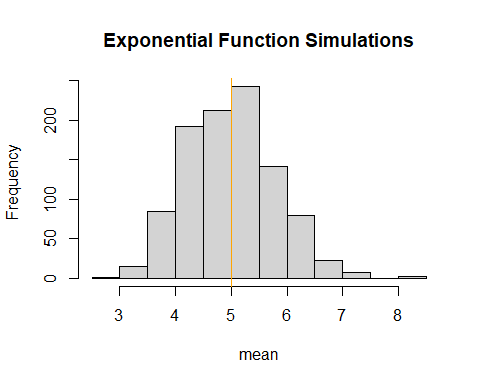
analytical\_mean <- mean(means\_exponentials)  
analytical\_mean

## [1] 4.993867

# analytical mean  
theory\_mean <- 1/lambda  
theory\_mean

## [1] 5

# visualization  
hist(means\_exponentials, xlab = "mean", main = "Exponential Function Simulations")  
abline(v = analytical\_mean, col = "red")  
abline(v = theory\_mean, col = "orange")



The analytics mean is 4.993867 the theoretical mean 5. The center of distribution of averages of 40 exponentials is very close to the theoretical center of the distribution.

## Question 2

Show how variable it is and compare it to the theoretical variance of the distribution..

# standard deviation of distribution  
standard\_deviation\_dist <- sd(means\_exponentials)  
standard\_deviation\_dist

## [1] 0.7931608

# standard deviation from analytical expression  
standard\_deviation\_theory <- (1/lambda)/sqrt(n)  
standard\_deviation\_theory

## [1] 0.7905694

# variance of distribution  
variance\_dist <- standard\_deviation\_dist^2  
variance\_dist

## [1] 0.6291041

# variance from analytical expression  
variance\_theory <- ((1/lambda)\*(1/sqrt(n)))^2  
variance\_theory

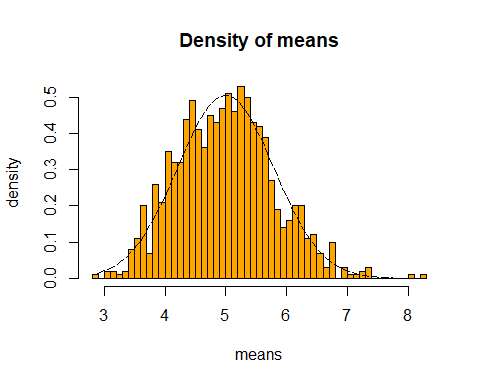
## [1] 0.625

Standard Deviation of the distribution is 0.7931608 with the theoretical SD calculated as 0.7905694. The Theoretical variance is calculated as ((1 / ??) \* (1/???n))2 = 0.625. The actual variance of the distribution is 0.6291041

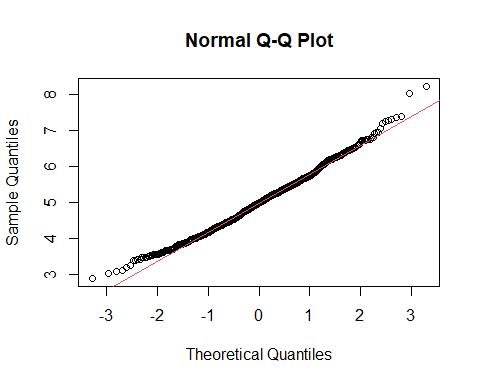
## Question 3

Show that the distribution is approximately normal.

xfit <- seq(min(means\_exponentials), max(means\_exponentials), length=100)  
yfit <- dnorm(xfit, mean=1/lambda, sd=(1/lambda/sqrt(n)))  
hist(means\_exponentials,breaks=n,prob=T,col="orange",xlab = "means",main="Density of means",ylab="density")  
lines(xfit, yfit, pch=22, col="black", lty=5)



# compare the distribution of averages of 40 exponentials to a normal distribution  
qqnorm(means\_exponentials)  
qqline(means\_exponentials, col = 2)



Due to Due to the central limit theorem (CLT), the distribution of averages of 40 exponentials is very close to a normal distribution.