Statistical Inference Project

Kieso Jan

8/21/2021

Instructions

This Report is created for:

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")
```

Assignment Task setup

```
# set seed for reproducability
set.seed(53)
# 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)
```

Task 1

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

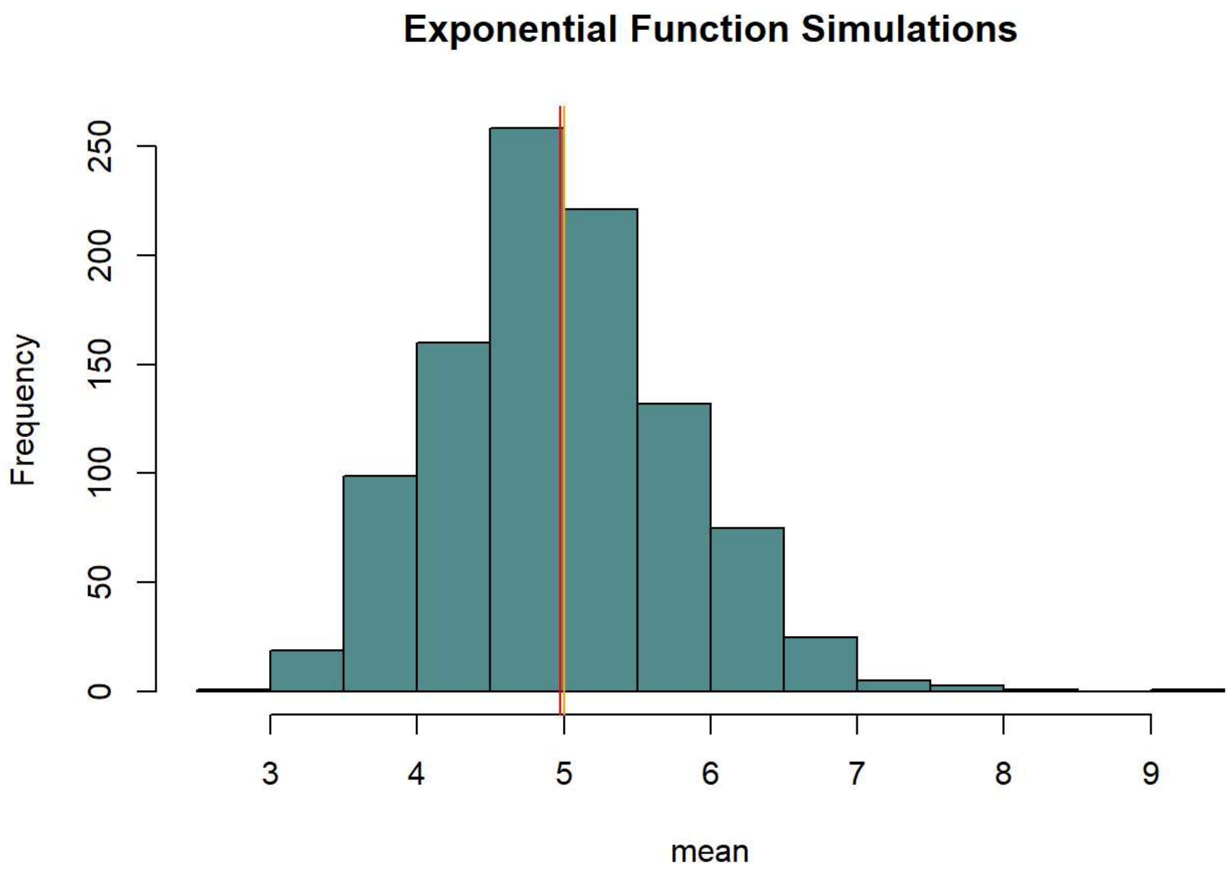
```
analytical_mean <- mean(means_exponentials)
analytical_mean
```

```
## [1] 4.973366
```

```
# Calculate analytical mean
theory_mean <- 1/lambda
theory_mean
```

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

```
# visualization
hist(means_exponentials, xlab = "mean", main = "Exponential Function Simulations", col = "darkslategray4")
abline(v = analytical_mean, col = "red")
abline(v = theory_mean, col = "orange")
```



Highlights of Task1:

1. The analytics mean is 4.973366 the theoretical mean 5.
2. The center of distribution of averages of 40 exponentials is very close to the theoretical center of the distribution.

Task 2

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

```
# Calculate standard deviation of distribution
SD_dist <- sd(means_exponentials)
SD_dist
```

```
## [1] 0.8149954
```

```
# Calculate standard deviation from analytical expression
SD_theory <- (1/lambda)/sqrt(n)
SD_theory
```

```
## [1] 0.7905694
```

```
# Calculate variance of distribution
variance_dist <- SD_dist^2
variance_dist
```

```
## [1] 0.6642175
```

```
# Calculate variance from analytical expression
variance_theory <- ((1/lambda)*(1/sqrt(n)))^2
variance_theory
```

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

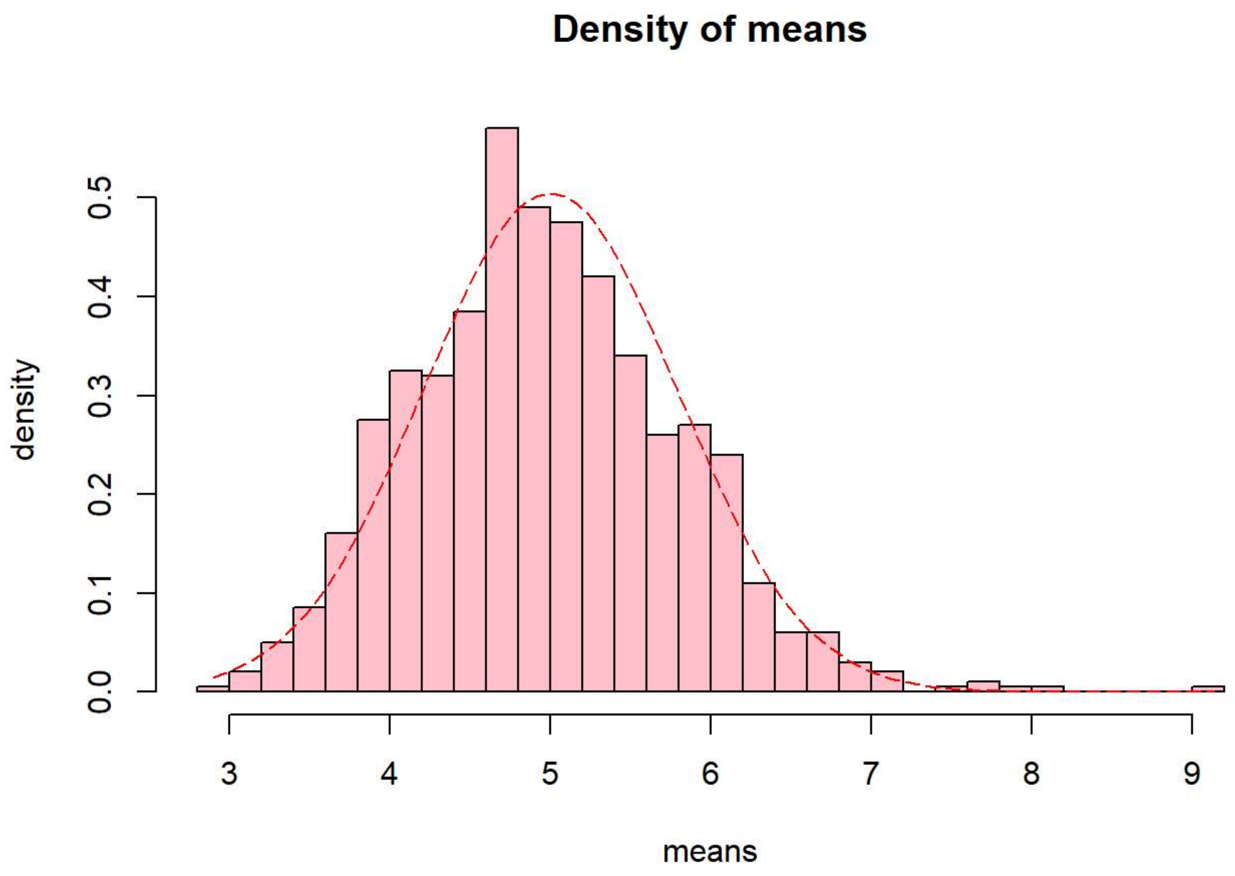
Highlights of Task2:

1. Standard Deviation of the distribution is 0.8149954 with the theoretical SD calculated as 0.7905694.
2. The Theoretical variance is calculated as $((1 / 0.2) * (1/\text{sqrt}(n))^2 = 0.625$. The actual variance of the distribution is 0.6642175

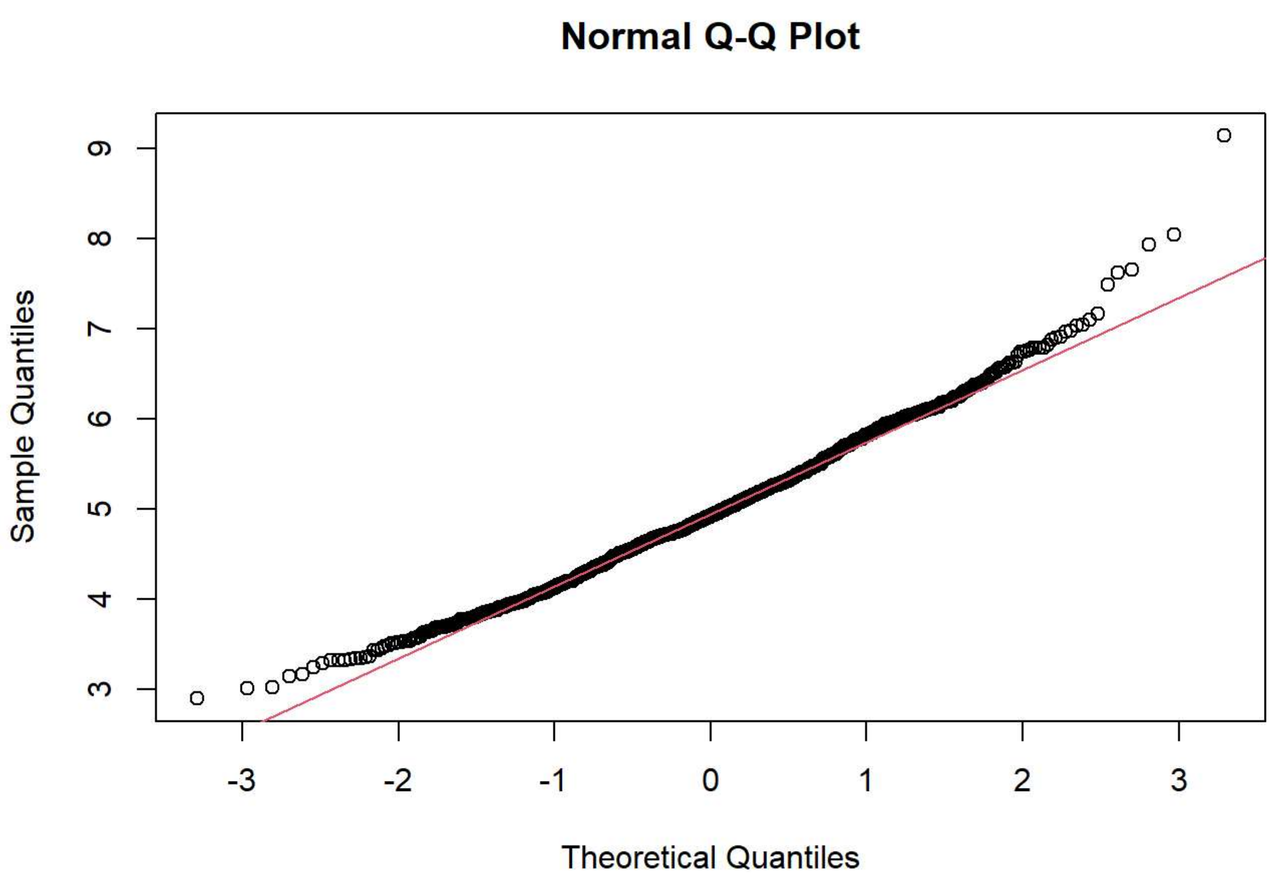
Task 3

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="pink",xlab = "means",main="Density of means",ylab="density")
lines(xfit, yfit, pch=22, col="red", lty=5)
```




```
# compare the distribution of averages of 40 exponentials to a normal distribution
qqnorm(means_exponentials)
qqline(means_exponentials, col = 2)
```



Highlights of Task3:

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