Statistical Inference project

Part1: Simulation of an Exponensial distribution

Description of the Problem

The exponential distribution can be simulated in R with rexp(n, lambda) where lambda is the rate parameter. The mean of exponential distributio n is 1/lambda and the standard deviation is also also 1/lambda. Set lambda = 0.2 for all of the simulations. In this simulation, you will investigate the distribution of averages of 40 exponential(0.2)s. Note that you will need to do a thousand or so simulated averages of 40 exponentials. Illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponential(0.2)s.

Questions 1 (show where the distribution is centered at and compare it to the theoretical center of the distribution) & Question 2 (show how variable it is and compare it to the theoretical variance of the distribution)

```
set.seed(3)
lambda <- 0.2
num_sim <- 1000
sample_size <- 40
sim <- matrix(rexp(num_sim*sample_size, rate=lambda), num_sim, sample_size)
row_means <-rowMeans(sim)
head(row_means)
```

```
## [1] 5.294484 3.266814 5.112187 4.602583 4.795966 3.983437
```

```
m=mean(row_means)
sd=sd(row_means)
```

Expected Standard Deviation

```
(1/lambda)/sqrt(40)
```

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

Variance of Simulation

```
var(row_means)
```

```
## [1] 0.6257575
```

Expected Variance

```
((1/lambda)/sqrt(40))^2
```

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

Solutions:

1.The expected center (5.0) is very close to the center of the distribution (4.9988).

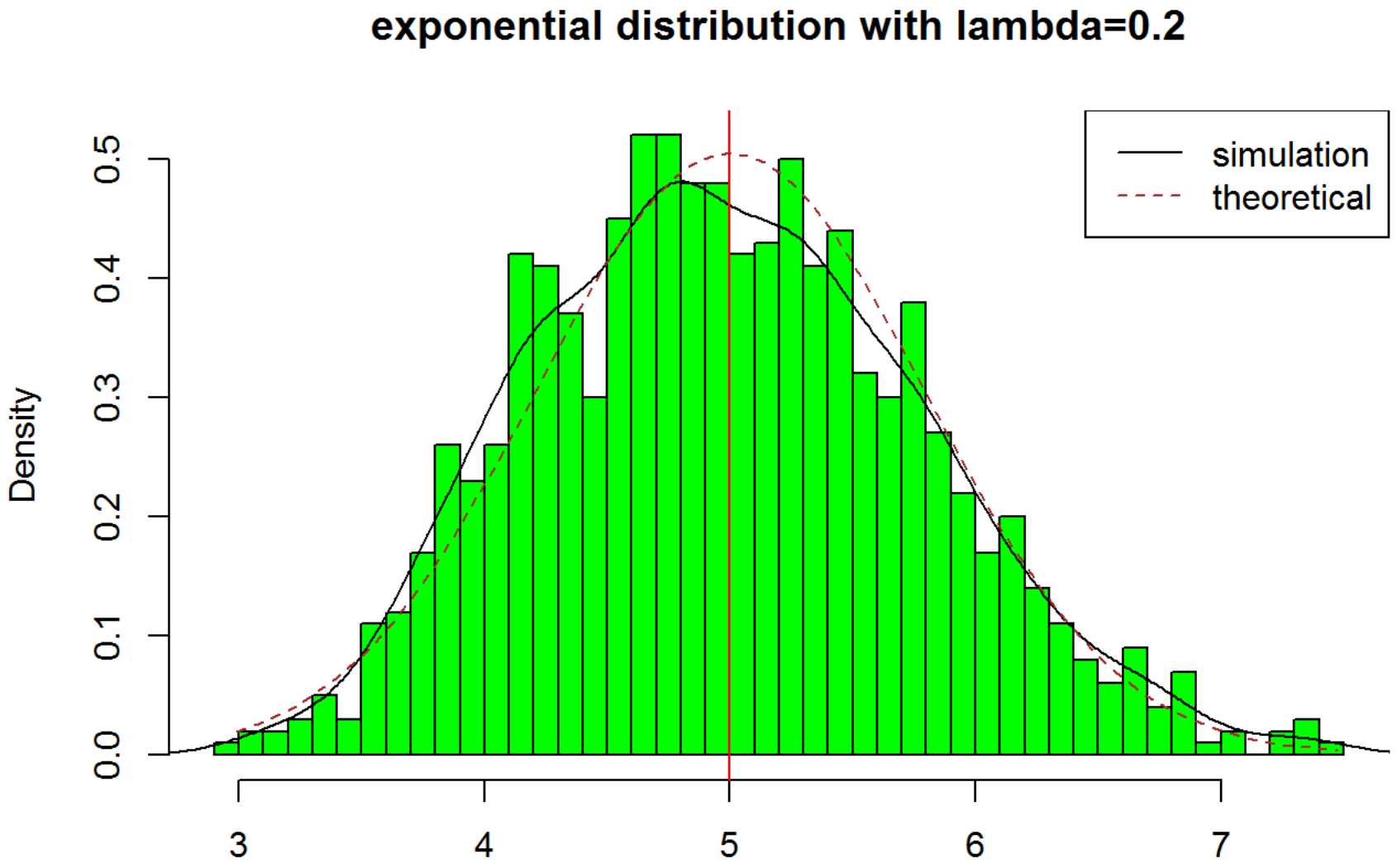
2.The standard deviation (0.7909) is also close to the expected standard deviation (0.79056).

Question 3 (show that the distribution is approximately normal)

```
library(ggplot2)
##ggplot(data=row_means,aes(x=x))+geom_histogram(aes(y=...density...),fill=I("olivegreen"##),ba
ndwidth=0.2,color=I("black"))+stat_function(fun=dnorm,arg=list(mean=5,sd=sd(row_means$rowMeans.
```

```
sim.)))

hist(row_means,breaks=50,prob=TRUE,main="exponential distribution with lambda=0.2",xlab="",col=
"green")
# density of the averages of samples
lines(density(row_means))
# theoretical center of distribution
abline(v=1/lambda, col="red")
# theoretical density of the averages of samples
xfit <- seq(min(row_means), max(row_means), length=100)
yfit <- dnorm(xfit, mean=1/lambda, sd=(1/lambda/sqrt(sample_size)))
lines(xfit, yfit, pch=22, col="brown", lty=2)
# add legend
legend('topright', c("simulation", "theoretical"), lty=c(1,2), col=c("black", "brown"))
```



Solution : 3:3.

The histogram plot depicts a distribution that is approximately normal (mean = 5; sd = .7909).

Question 4 (evaluate the coverage of the confidence interval for 1/lambda: $\bar{X} \pm 1.96 Sn$???)

```
mean(row_means) + c(-1, 1) * 1.96 * sd(row_means)/sqrt(1000)
```

```
## [1] 4.937590 5.035649
```

Solution: 4. The 95% confidence interval for the mean of the means is 4.950 - 5.047.