Statistical Inference Project Pt 1

Hsin Chih Chen

10/6/2021

Instructions

1. Demonstrate sample mean and compare to theoretical mean of distribution
2. Indicate sample's variance and compare to theoretical variance of the distribution.
3. Show that the distribution is approximately normal.

Load Applicable Libraries

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

Preliminary Setup for the random exponential generation

```
# set seed for random number consistency
set.seed(69)

# set lambda to 0.2
lambda <- 0.2

# assign samples (n) and simulation (sims) are allocated
n <- 40
sims <- 1000

# simulate the setup for repeated random distribution
sim_exp <- replicate(sims, rexp(n, lambda))

# calculate mean of exponential
means_exp <- apply(sim_exp, 2, mean)
```

Section 1

Demonstrate distribution's center line for experimental and theoretical mean

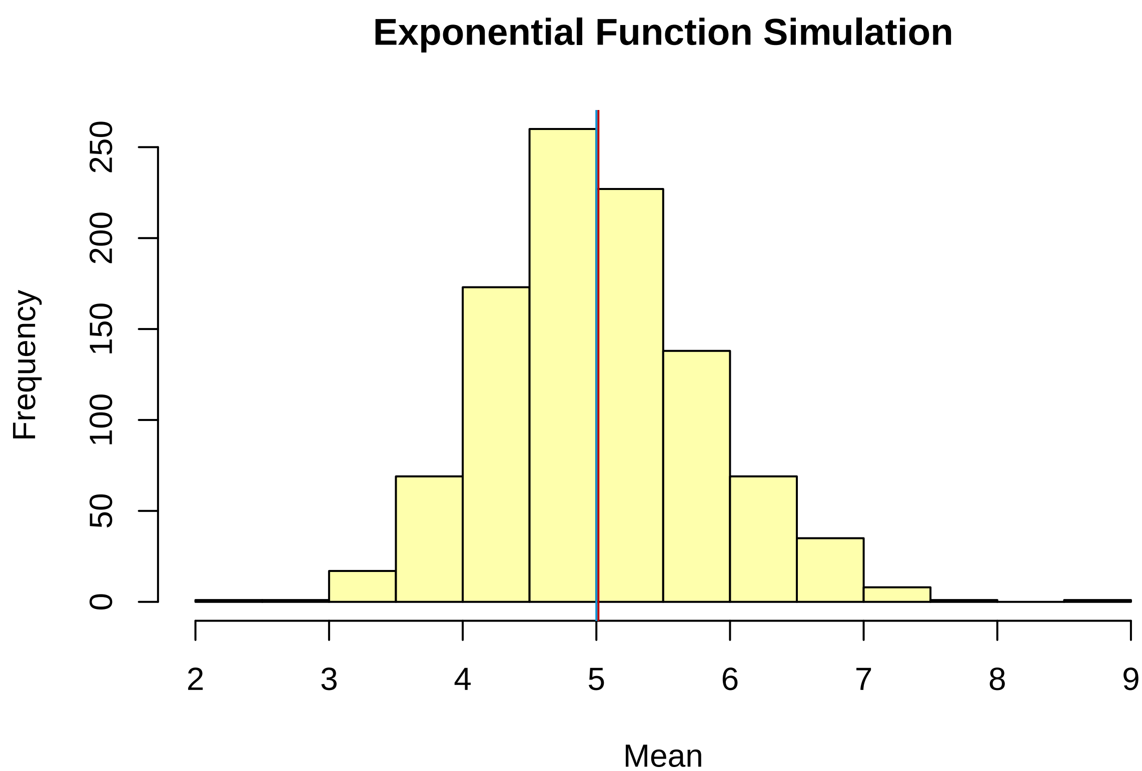
```
# calculate mean of the experimental simulation
e_mean <- mean(means_exp)
e_mean
```

```
## [1] 5.015158
```

```r
# calculate mean from theoretical expression
t_mean <- 1/lambda
t_mean
```

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

```r
# Generate the histogram for the experimental mean simulation while add the averages of
hist(means_exp, xlab = "Mean", main = "Exponential Function Simulation", col = "#FEFFAC")
abline(v = e_mean, col = "#BB0000")
abline(v = t_mean, col = "#1191D3")
```



The experimental mean is 5.015158 and the theoretical mean 5. The center of distribution of averages of 40 exponentials is very close to the theoretical center of the distribution.

Section 2

Demonstrate distribution's variance for experimental and theoretical distribution.

```r
# standard deviation of the exponential simulation
stdev_e <- sd(means_exp)
stdev_e
```

```
## [1] 0.7982937
```

```r
# standard deviation from theoretical expression
stdev_t <- (1/lambda)/sqrt(n)
stdev_t
```

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

```r
# variance of distribution
var_dist <- stdev_e^2
var_dist
```

```
## [1] 0.6372728
```

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

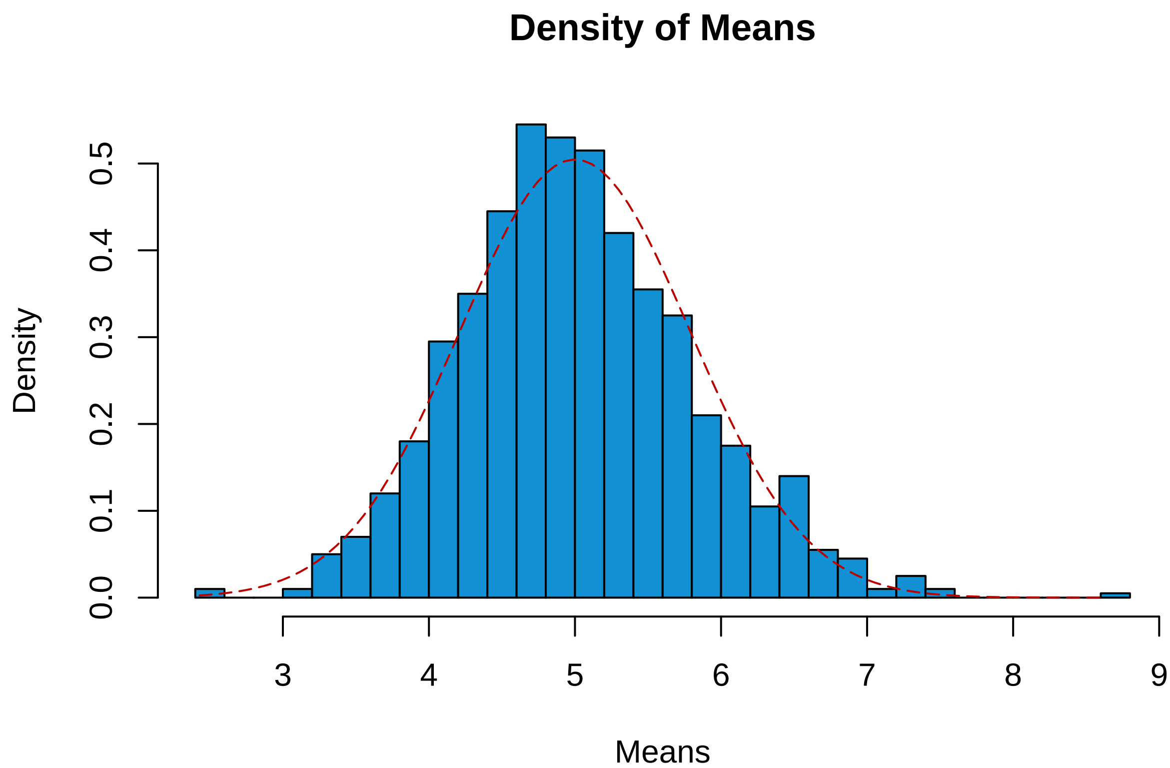
```
## [1] 0.625
```

Experimental stdev is 0.7982937 with the theoretical stdev calculated as 0.7905694. The Theoretical variance is calculated as ((1 / lambda) * (1/sqrt(n))^2 = 0.625 while the experimental variance is 0.6372728

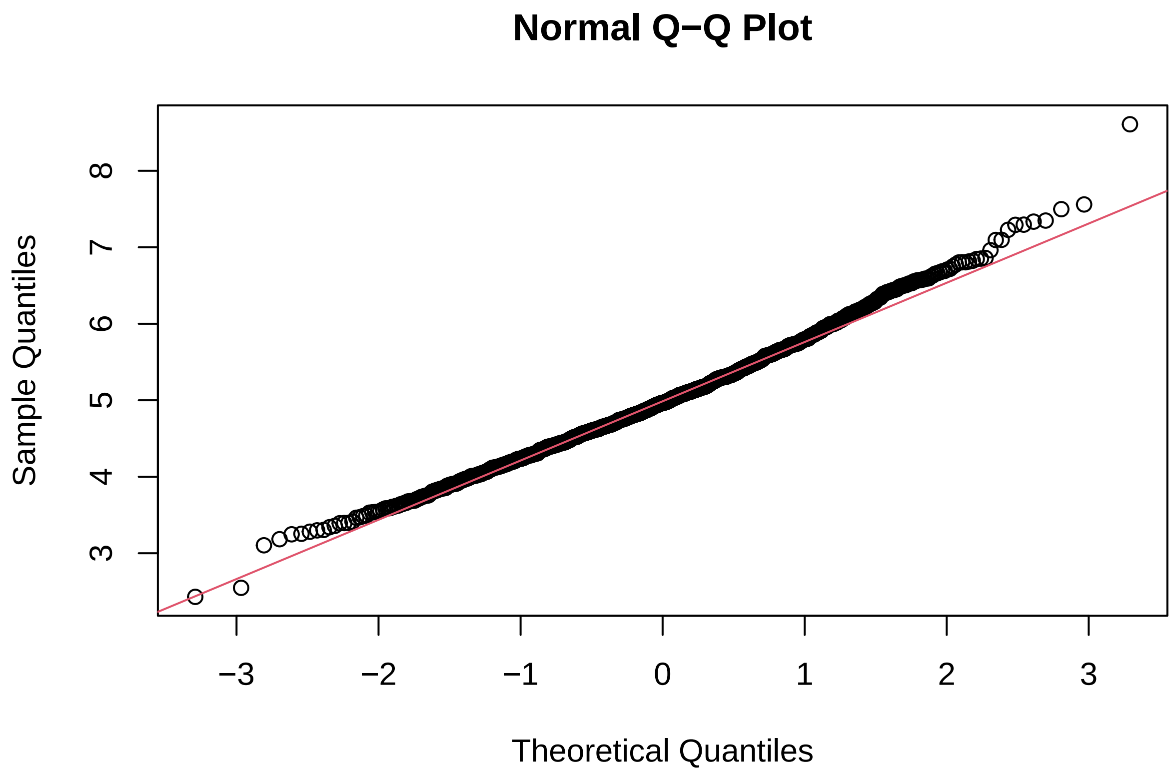
Section 3

Indicate the distribution is approximately normal.

```r
x_set <- seq(min(means_exp), max(means_exp), length=100)
y_set <- dnorm(x_set, mean=1/lambda, sd=(1/lambda/sqrt(n)))
hist(means_exp,breaks=n,prob=T,col="#1191D3",xlab = "Means",main="Density of Means",ylab="Density")
lines(x_set, y_set, pch=22, col="#BB0000", lty=5)
```




```r
# Compare the distribution of averages of 40 exponentials to a normal distribution
qqnorm(means_exp)
qqline(means_exp, col = 2)
```



Based on Central Limit Theorem (CLT), the distribution of 40 exponentials averages is very close to a normal distribution.