

# A Simulation Exercise

Vijayeshwari Palakurthi

7/4/2020

The exponential distribution can be simulated in R with `rexp(n, lambda)` where `lambda` is the rate parameter. The mean of exponential distribution is  $1/\lambda$  and the standard deviation is also  $1/\lambda$ .

```
set.seed(1)
lambda <- 0.2 # Set lambda = 0.2 for all of the simulations.
n <- 40       # In this simulation, we investigate the distribution of averages
              # of 40 exponentials.
simulations <- 1:1000 # We need to do a thousand or so simulated averages
averages <- sapply(simulations, function(x) { mean(rexp(n, lambda)) })
```

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

When we calculate sample and theoretical mean, we see that both lie close together.

```
mean(averages)
```

```
## [1] 4.990025
```

```
1/lambda
```

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

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

From the CLT we know that  $\bar{X}$  approximately follows  $N(\mu, \sigma^2/n)$ . We know  $\sigma$  to be  $1/\lambda$ . As such it follows that the theoretical standard deviation is:

```
(1/lambda)/sqrt(40) # Theoretical standard deviation
```

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

```
sd(averages)      # actual standard deviation
```

```
## [1] 0.7817394
```

```
# And the variances
((1/lambda)/sqrt(40))^2
```

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

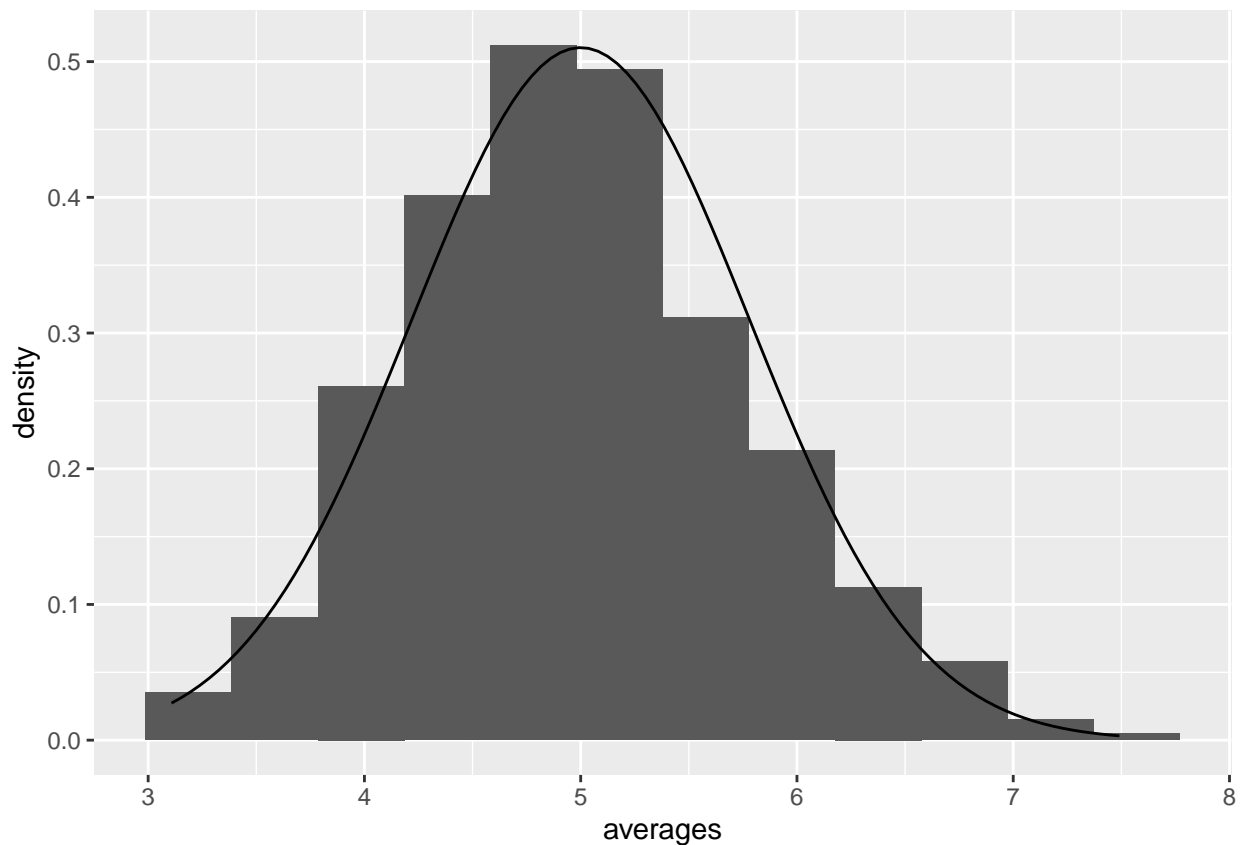
```
sd(averages)^2
```

```
## [1] 0.6111165
```

### 3. Show that the distribution is approximately normal.

To do so, we plot an histogram of the sampled means and overlay the normal distribution with mean 5 and standard deviation 0.7817394 on top of it. We see that the normal distribution indeed closely matches the barplot of the means.

```
library(ggplot2)
# Sturges' formula
k <- ceiling(log2(length(simulations)) + 1)
bw <- (range(averages)[2] - range(averages)[1]) / k
averages.sd <- sd(averages)
p <- ggplot(data.frame(averages), aes(x=averages))
p <- p + geom_histogram(aes(y=..density..), binwidth=bw)
p <- p + stat_function(fun = dnorm, args=list(mean=5, sd=averages.sd))
p
```



#### 4. Evaluate the coverage.

Evaluate the coverage of the confidence interval for  $1/\lambda$ :

$$\bar{X} \pm 1.96 \frac{S}{\sqrt{n}}$$

```
mean(averages) + c(-1,1) * 1.96 * sd(averages) / sqrt(length(averages))
```

```
## [1] 4.941572 5.038478
```