Title: Statistical Inference Course Project: Part 1

Author: David Brown

This report will display an exponential distribution. To demonstrate how the Central Limit Theorem works, I will collect the average of 40 exponentials 1000 times. The sample mean will be an unbiased estimate of the population mean, and the sample distribution will be Gaussian.

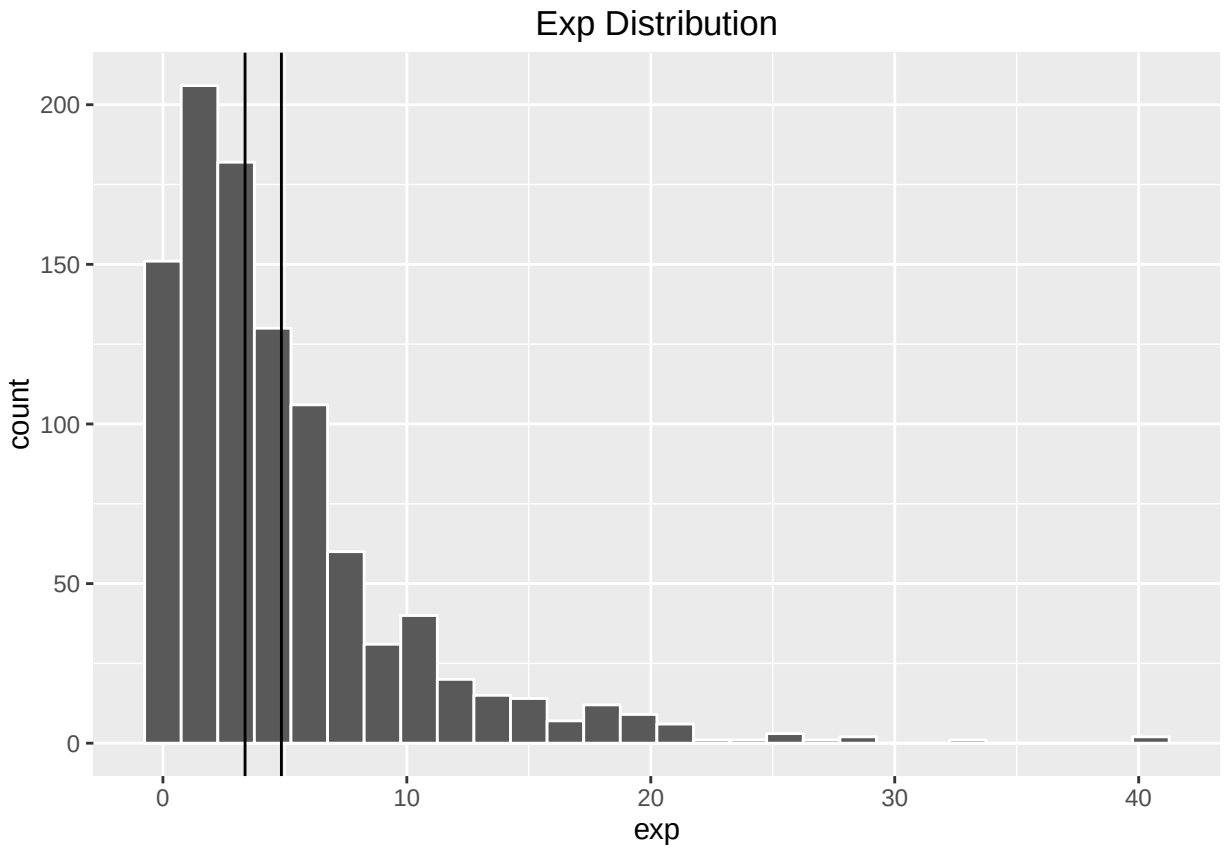
Using rexp I generated 1000 exponentials with lambda = 0.2. The average of this distribution is roughly 1/0.2; additionally, the standard deviation is also roughly 1/0.2. To demonstrate the Central Limit Theorem, I sampled 40 exponentials from the population 1000 times. I used a for loop and the sample function.

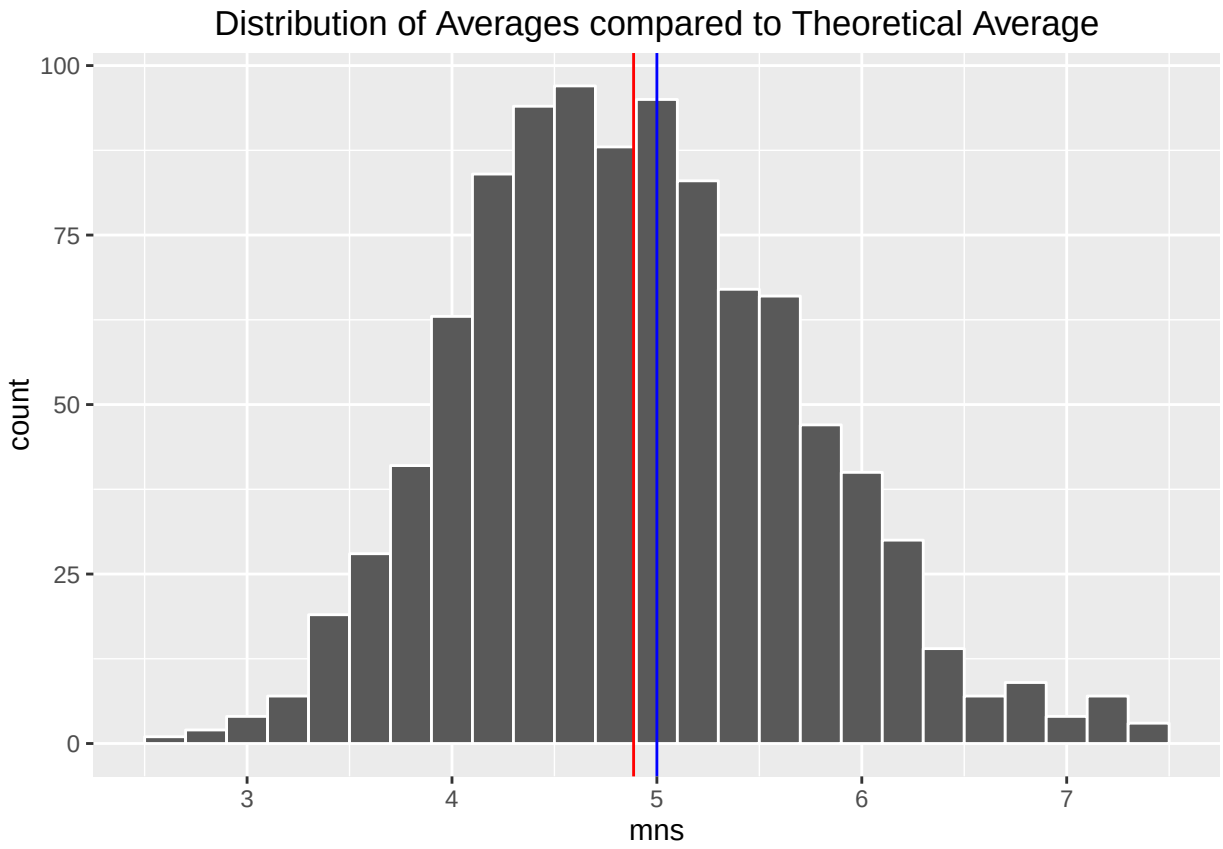
```
set.seed(100)

exp <- rexp(1000, 0.2)
```

```
mns = NULL

for (i in 1:1000) mns = c(mns, mean(sample(exp, 40, replace = TRUE)))
```





Averages

```
##    theoretical    pop_average sample_average
##       5.000000       4.857406       4.886417
```

Variances

```
## theoretical  pop_distro  sample_var
##    0.025000   25.624002    0.025624
```

The sample mean is a good, unbiased estimator of the population mean. The population variance is the same as the population mean. The sample distribution is Gaussian.

ggplot() + geom_histogram(aes(exp), color = "white", binwidth = 1.5) + geom_vline(xintercept = c(median(exp), mean(exp))) + ggtitle("Exp Distribution")

ggplot() + geom_histogram(aes(mns), color = "white", binwidth = .2) + geom_vline(xintercept = mean(mns), color = "red") + geom_vline(xintercept = 1/0.2, color = "blue") + ggtitle("Distribution of Averages vs Theoretical Distribution")