

## Statistical Inference Course Project 1

### Introduction

In this project you will investigate the exponential distribution in R and compare it with the Central Limit Theorem. 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  $\lambda = 0.2$  for all of the simulations. You will investigate the distribution of averages of 40 exponentials. Note that you will need to do a thousand simulations.

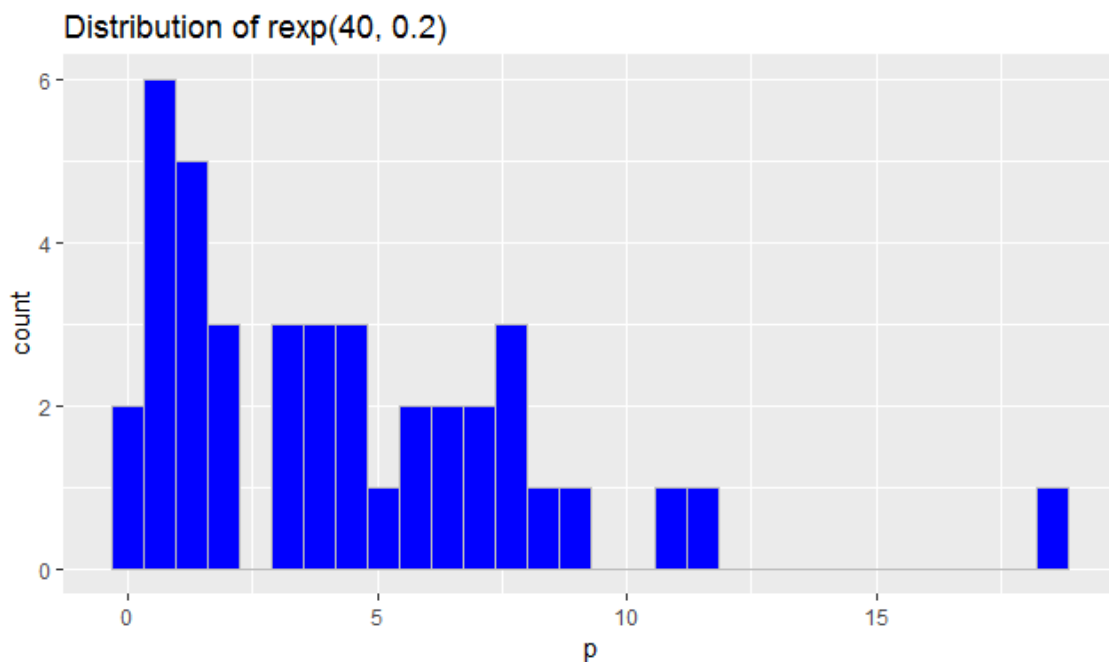
```
library(ggplot2)

p <- rexp(40, 0.2)

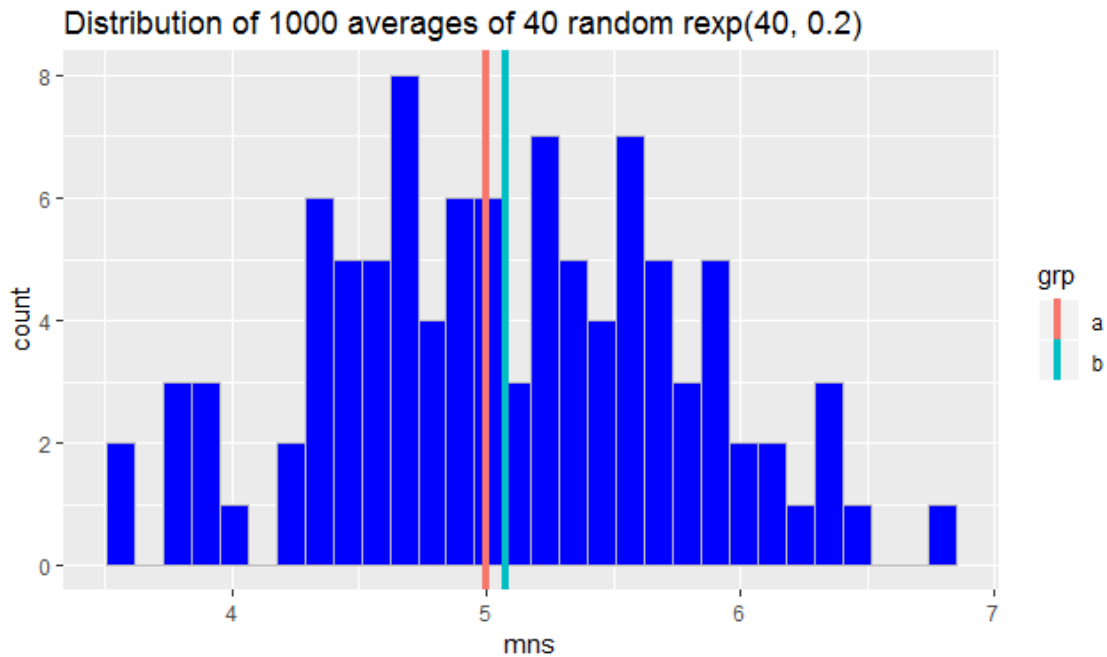
g = ggplot(data.frame(p = p), aes(x = p))

g = g + geom_histogram(color = "gray", fill = "blue") + labs(title = "Distribution of rexp(40, 0.2)")

print(g);
```



- From the above graph, we can see that the distribution is similar to an exponential function. Next graph is the distribution of 1000 averages of 40 random `rexp(40, 0.2)`.



In the above plot the red line shows the theoretical center of the distribution which is 5 ( $1 / 0.2$ ). The blue line shows the center of the distribution. It is evident that both are very close. The actual mean of the distribution is shown below

```
print(mean(mns));
##5.026728
```

- The variance of the 1000 averages of 40 random rexp(40, 0.2) is shown as below.

```
print(var(mns))
##0.5072297
```

- The theoretical variance of the distribution should be  $(1/\lambda)^2/n$  which is shown as below

```
(1/0.2)^2/40
##0.625
```

- The mean and variance of the 1000 averages of 40 random rexp(40, 0.2), and the plot, confirm this distribution is approximately normal.