Simulation Exercise Report

Borye

Friday, November 07, 2014

Introduction

This is a project for the statistical inference class hold by coursera. In this part of the project, I will use simulation to explore inference and do some simple inferential data analysis.

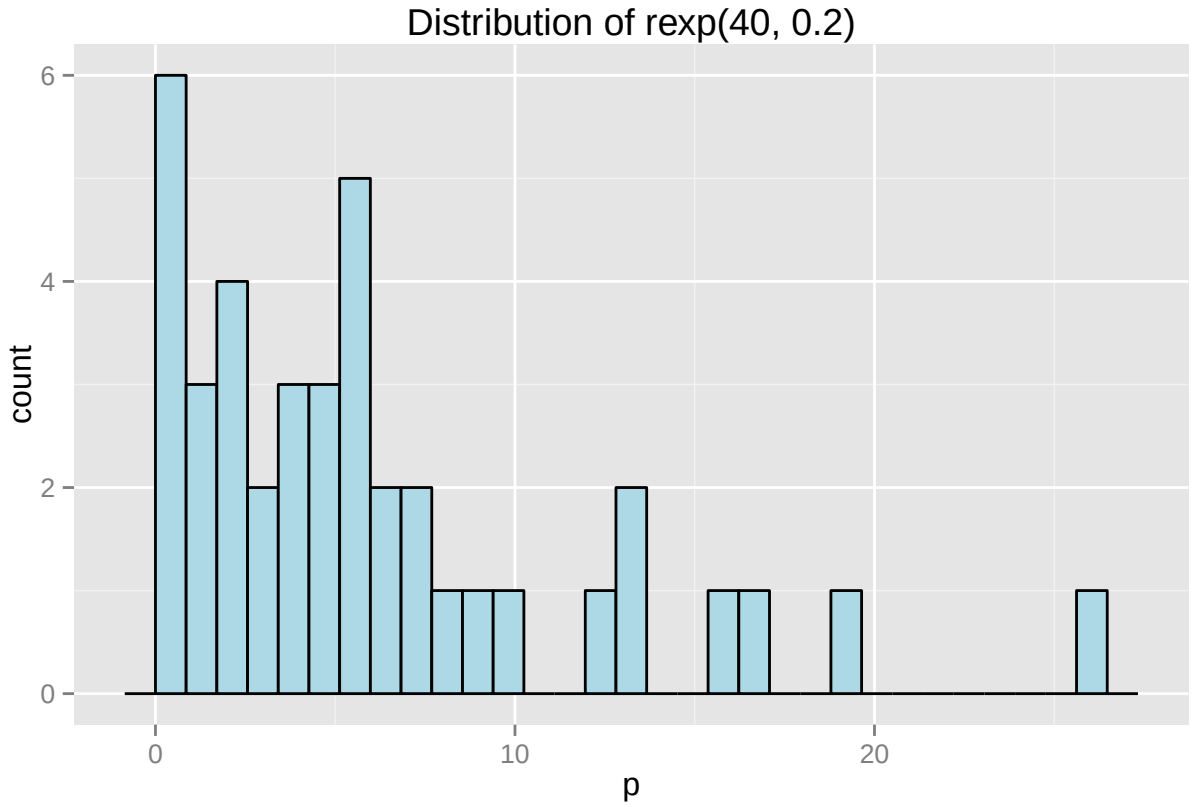
The exponential distribution can be simulated in R with rexp(n, lambda) where lambda is the rate parameter. The mean of exponential distribution is 1/lambda and the standard deviation is also also 1/lambda. **Set lambda = 0.2** for all of the simulations. In this simulation, you will investigate the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

Question 1, Show where the distribution is centered at and compare it to the theoretical center of the distribution

First, I plot the original distribution of the rexp(40, 0.2)

```
library(ggplot2)
p <- rexp(40, 0.2)
g = ggplot(data.frame(p = p), aes(x = p))
g = g + geom_histogram(color = "black", fill = "lightblue") + labs(title = "Distribution of rexp(40, 0.
g
```

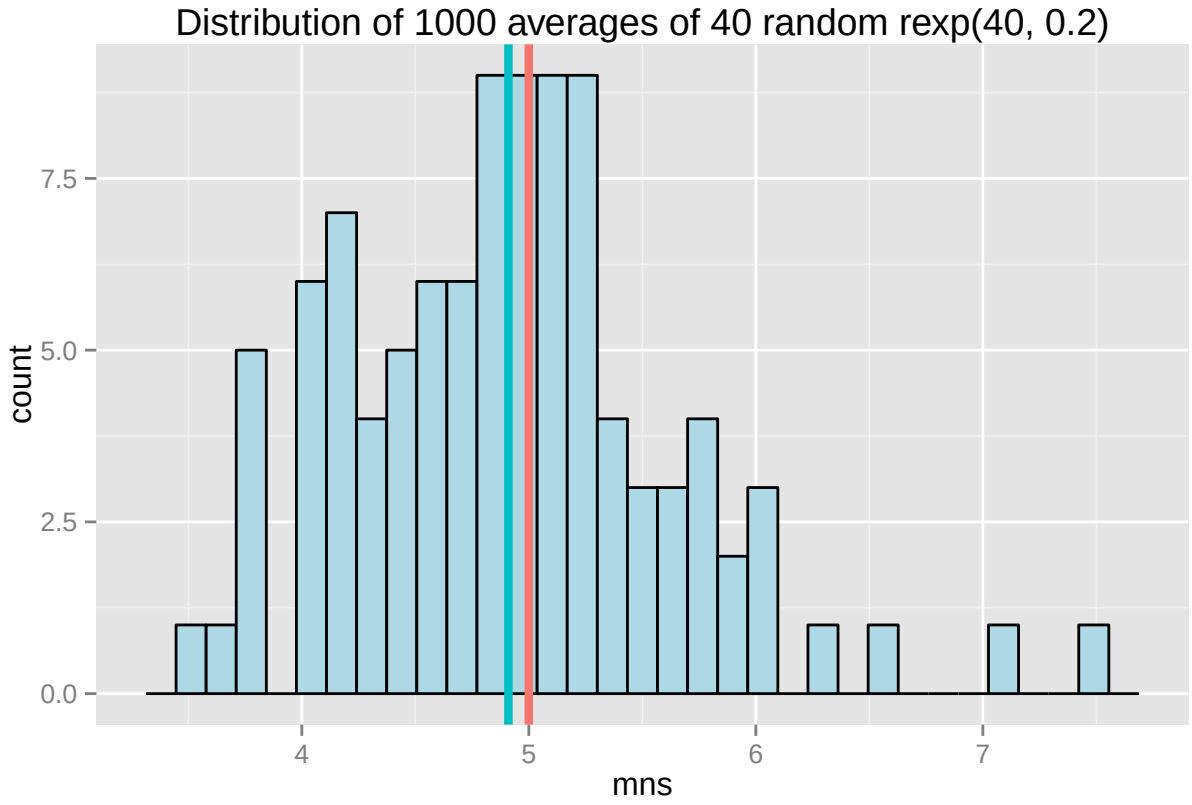
```
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
```



We could seen from this plot, that the distribution is similar as a exponential function. Next plot is the distribution of 1000 averages of 40 random rexp(40, 0.2)

```
mns = NULL
for (i in 1 : 100) mns = c(mns, mean(rexp(40, 0.2)))
vlines <- data.frame(xint = c(5, mean(mns)), grp = c("a", "b"))
g = ggplot(data.frame(mns), aes(x = mns))
g = g + geom_histogram(color = "black", fill = "lightblue") + geom_vline(data = vlines, aes(xintercept =
g
```

```
## stat_bin: binwidth defaulted to range/30. Use 'binwidth = x' to adjust this.
```



We could seen from this plot that the **red line shows the theoretical center of the distribution which is 5 (1/0.2)**. While the **blue line shows the center of the distribution**. We can seen from this plot that these two lines are very close to each other. The actual mean of the distributio is shown below.

```
mean(mns)
```

```
## [1] 4.910066
```

Question 2, show how variable it is and compare it to the theoretical variance of the distribution

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

```
var(mns)
```

```
## [1] 0.5237364
```

The theoretical variance of the distribution should be **(1/lambda)^2/n** which shownd as below

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

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

Question 3, show that the distribution is approximately normal

From the plot, and the mean and varience of the 1000 averages of 40 random rexp(40, 0,2), we could believe that this distribution is approximately normal.