Exploring Properties of the Exponential Distribution

Bosah Chude

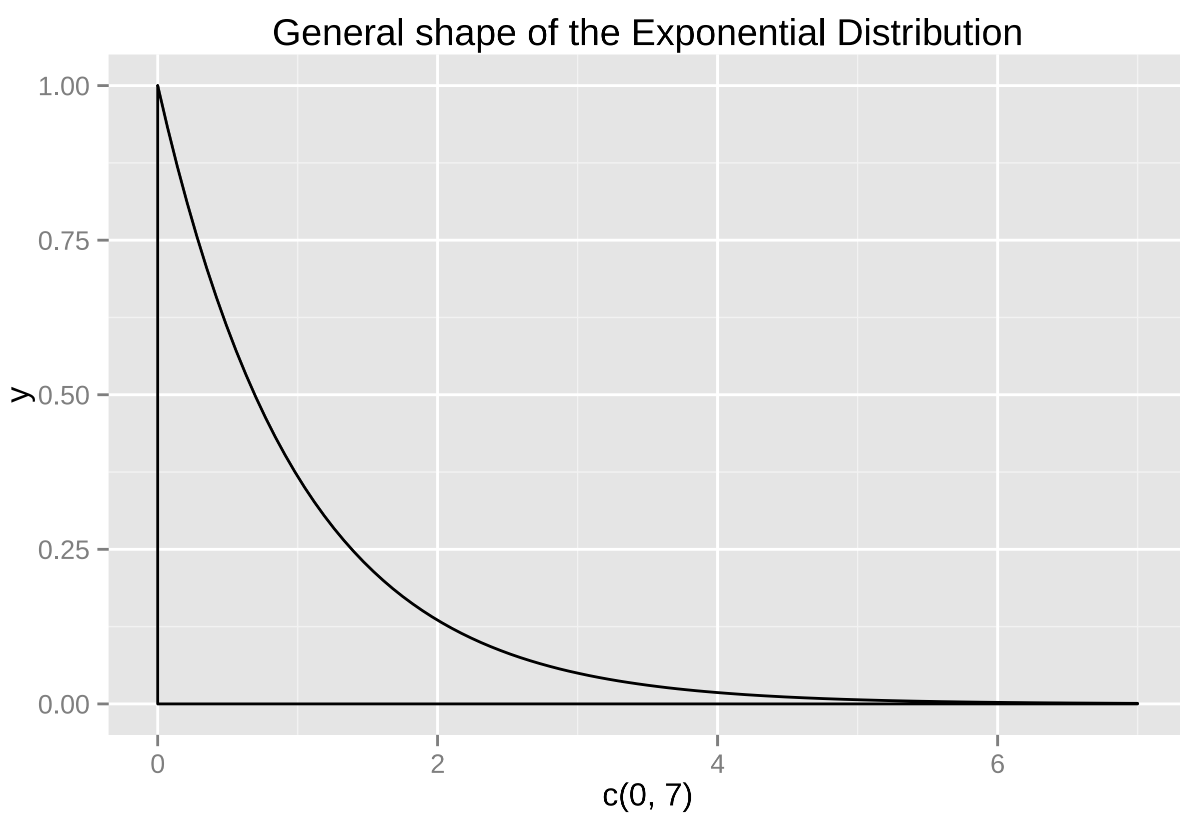
Thursday, November 20, 2014

Solution to Statistical Inference course project (Section 1).

I would start by displaying the general density plot of the exponential distribution

```
#Load library and set seed to ensure reproducibility
library(ggplot2)
library(gridExtra)
set.seed(2334)

#Standard Exponential plot
qplot(c(0, 7), stat = "function", fun = dexp,
      geom = "density", main = "General shape of the Exponential Distribution")
```



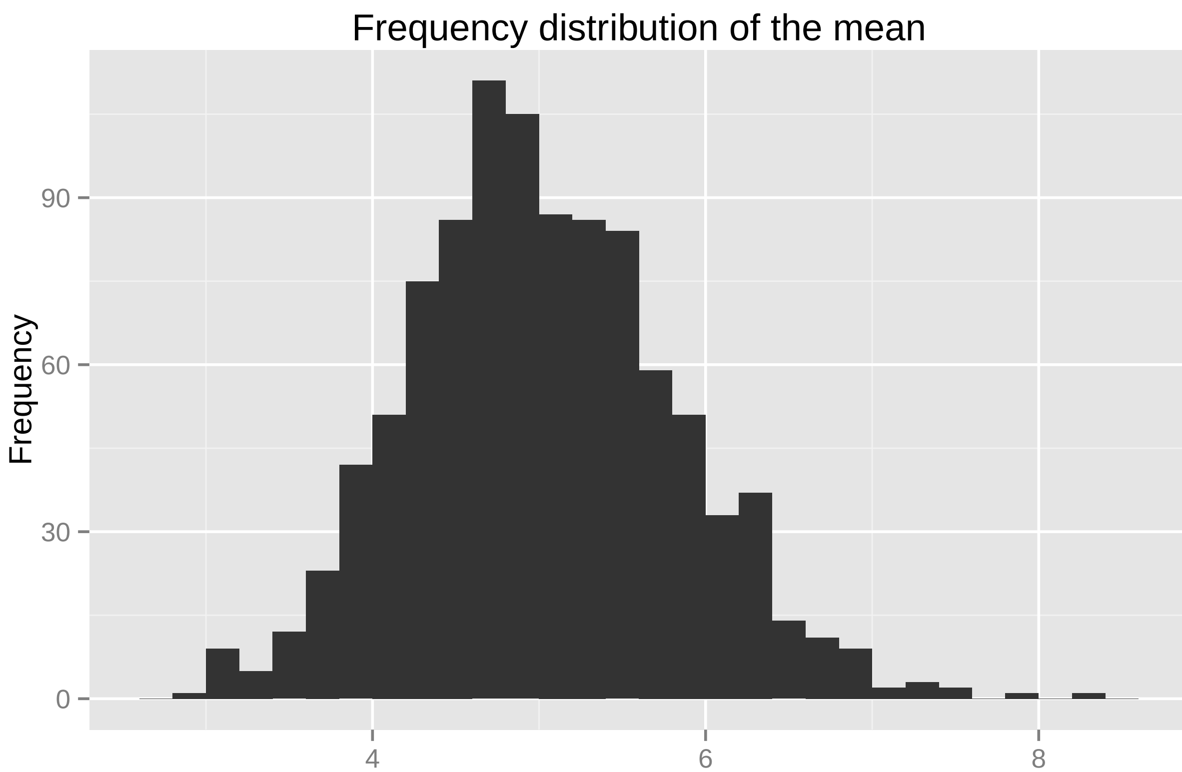
Q1. Show where the distribution is centered at and compare it to the theoretical center of the distribution.

The center of the distribution is its mean.

This can be gotten by running 1000 simulation and taking their mean.

```r
#Running 1000 simulation and taking the average.
distMean = NULL
for (i in 1 : 1000) distMean = c(distMean, mean(rexp(40, 0.2)))

#Visual plot of the distribution
qplot(distMean, geom = "histogram", binwidth = 0.2,
      ylab = "Frequency", xlab = "", main = "Frequency distribution of the mean")
```




```r
#Determinig the center (i.e mean)
mean(distMean)
```

```
## [1] 5.019324
```

```r
#The theoritical center is 1/lambda.
#lambda is 0.2 for our simulation
1/0.2
```

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

From the above computations, the distribution mean is **5.0193244** and the theoretical mean is **5**.

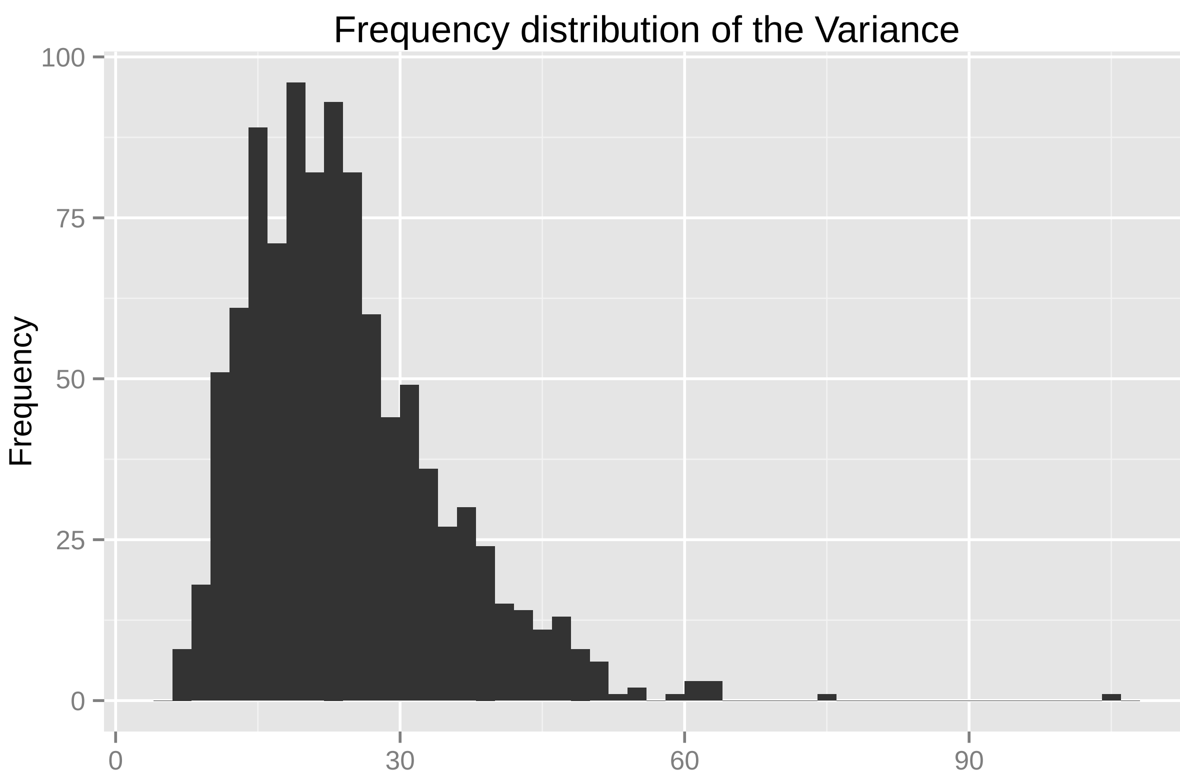
They are almost identical.

Q2. Show how variable it is and compare it to the theoretical variance of the distribution.

Computing the distribution variance of 1000 simulations.

```
#Running 1000 simulation and taking thier variance.
distVar = NULL
for (i in 1 : 1000) distVar = c(distVar, var(rexp(40, 0.2)))

#plot of variance
qplot(distVar, geom = "histogram", binwidth = 2,
      ylab = "Frequency", xlab = "", main = "Frequency distribution of the Variance")
```




```
#Center of the variance. (Average variance of the distribution)
mean(distVar)
```

```
## [1] 24.21152
```

```
#The theoretical variance is the square of the standard deviation.
#The standard deviation is 1/lambda
(1/0.2)^2
```

```
## [1] 25
```

From the following, the distribution variance is **24.2115215**. Which is very close to the theoretical variance of **25**.

Q3. Show that the distribution is approximately normal.

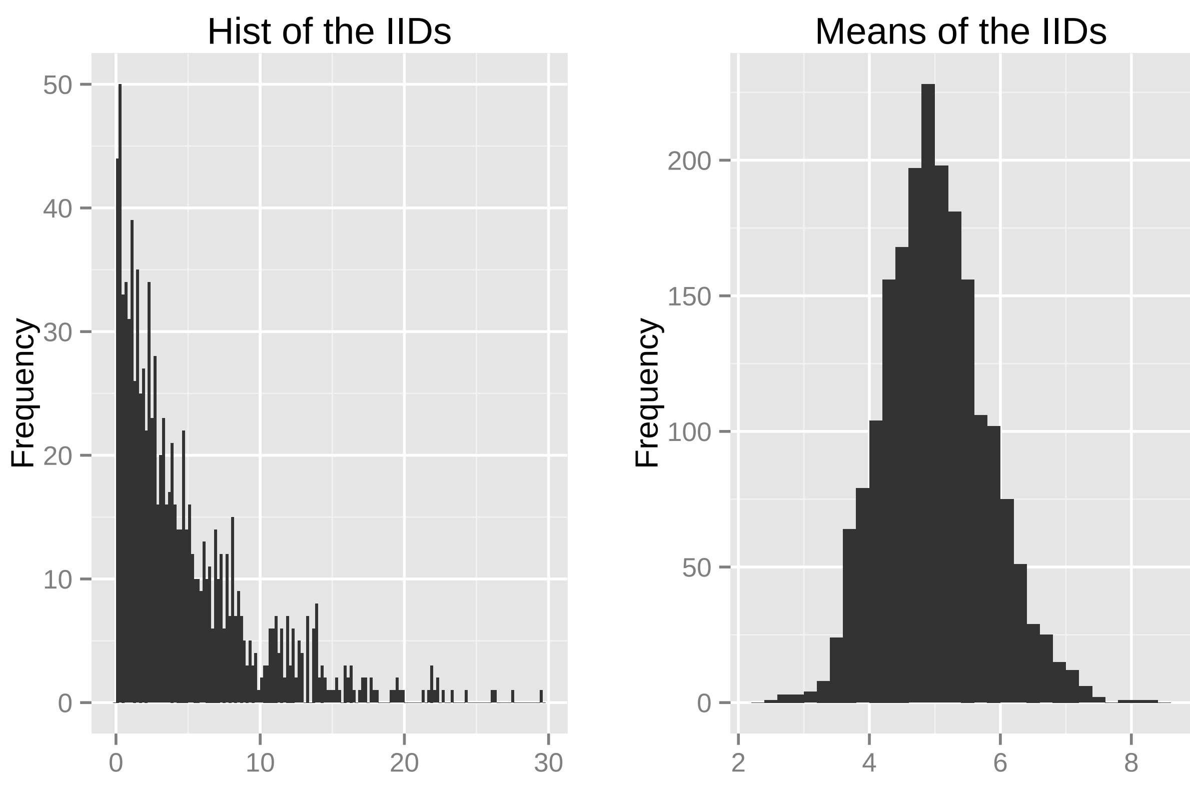
The distribution should be approximately normal due to the central limit theorem which states that the distribution of the sum (or average) of a large number of independent, identically distributed variables will be approximately normal, regardless of the underlying distribution.

We can test this as follows.

```
#Simulating 1000 random exponentials
firstSim <- rexp(1000, 0.2)
firstPlot <- qplot(firstSim, geom = 'histogram', binwidth = 0.2,
                   ylab = "Frequency", xlab = "", main = "Hist of the IIDs")

#Simulating the mean of 2000 random exponentials
secondSim = NULL
for (i in 1 : 2000) secondSim = c(secondSim, mean(rexp(40, 0.2)))
secondPlot <- qplot(secondSim, geom = 'histogram',  binwidth = 0.2,
                    ylab = "Frequency", xlab = "", main = "Means of the IIDs")

grid.arrange(firstPlot, secondPlot, ncol = 2)
```



From the above chart, the distribution of the mean resembles the normal distribution as predicted by the central limit theorem.