Sample Means of an Exponential Distribution

Brian Baquiran

November 22, 2014

The exponential distribution represents the distribution of waiting times in a Poisson process with rate of change λ. The probability density function for the exponential distribution is

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

The distribution is supported on the interval $[0, \infty)$. The mean of an exponential distribution is $\frac{1}{\lambda}$ and the standard deviation is also also $\frac{1}{\lambda}$.

In R the exponential distribution is simulated with rexp(n, lambda) where lambda is the rate parameter.

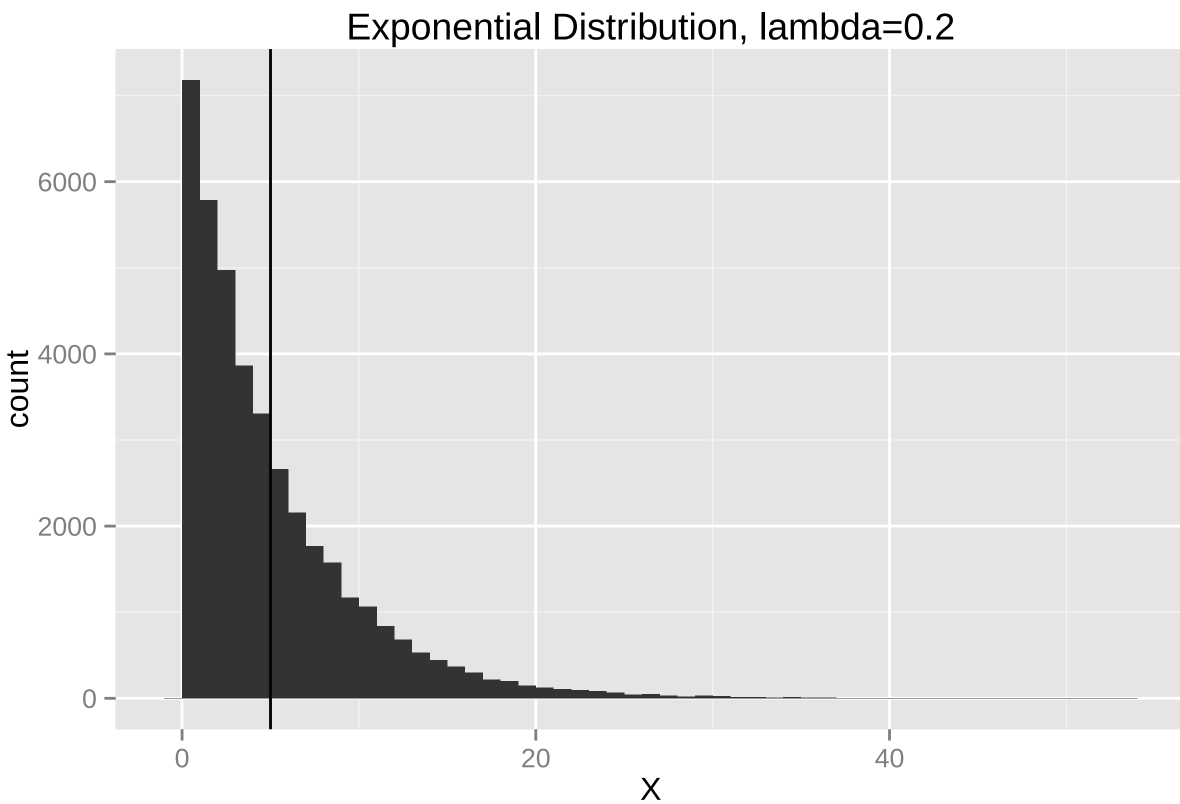
```
n <- 40
lambda <- 0.2

# our expected mean and standard deviation
mu <- 1/lambda
sigma <- 1/lambda
```

We expect to observe a mean $\mu = 5$ and standard deviation of $\sigma = 5$ from our simulation.

Simulation

To get a more intuitive understanding of the exponential distribution, let us first see what it looks like. We ask R to give us 40,000 random values from the distribution, and plot them on a histogram. The distribution mean of 5 is indicated by the vertical line.

```
library("ggplot2")
# 1,000 simulations
num.sims <- 1000
X <- rexp(n*num.sims,lambda)
p<-qplot(X,geom=c("histogram"),binwidth=1, main="Exponential Distribution, lambda=0.2") + geom_vline(ae
print(p)
```



Mean and Standard Deviation

We compute the mean and standard deviation of the random values and compare them to our expected value and standard deviation of $\frac{1}{\lambda}$.

```
print(Xbar<-mean(X))
```

```
## [1] 5.052
```

The mean for the 40,000 generated random values is 5.0523.

```
sd(X)
```

```
## [1] 5.037
```

The standard deviation for the random values is 5.0374.

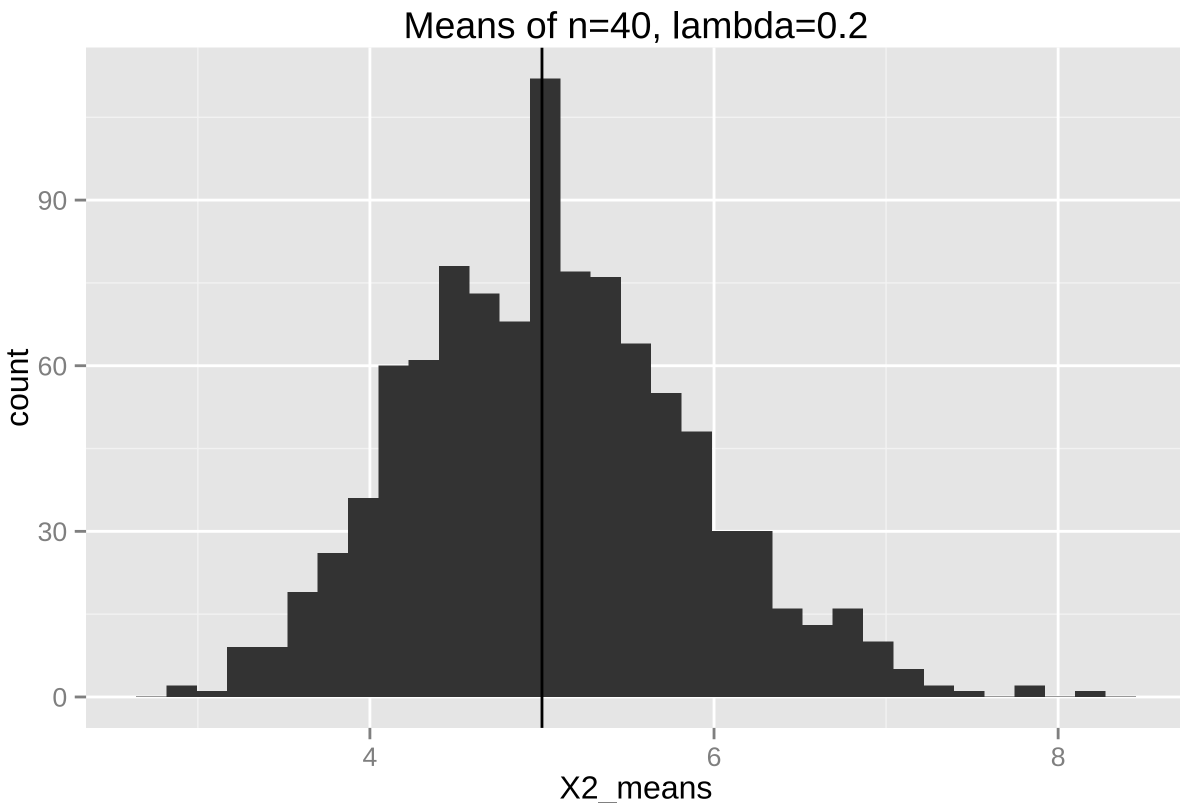
Both the observed mean and standard deviation are close to $\frac{1}{\lambda}$ which we know to be 5.

Sample means for n=40.

We divide the 40,000 random values we got earlier into 1000 samples of n=40, and get the mean of each sample. The expected value is indicated by a vertical line.

```
X2 <- matrix(X,nrow=num.sims)
X2_means <- apply(X2,1,mean)
#hist(X_means,breaks=50, main ="Means of sample size 40")
p<-qplot(X2_means,geom=c("histogram"),
        main="Means of n=40, lambda=0.2") +
        geom_vline(aes(xintercept=1/lambda))
print(p)
```

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



The distribution of the sample means *looks* normal and is centered around the distribution mean of 5. The mean of the sample means is in fact 5.0523.

Standard error for n=40.

The standard error of means of size n is $\frac{\sigma}{\sqrt{n}}$.

```
std.err = sigma/sqrt(n)
```

We find that for samples of size n=40, standard error is 0.7906. We can compare this to the standard deviation of the sample means.

```
sd(X2_means)
```

```
## [1] 0.823
```

So this confirms our understanding that the standard deviation of the sample means is estimated by the standard error $\frac{\sigma}{\sqrt{n}}$.