Central Limit Theorem in Exponential Distributions

Connie Wang

June 1, 2018

Overview

This report demonstrates the Central Limit Theorem using randomly generated exponential distributions. We show that the sample mean and variance of an exponential distribution with lambda (rate) = 0.2 approaches the theoretical mean and variance (mean = standard deviation = 1/lambda = 5) as the sample size increases. We also show that the distribution of the means of 40 exponentials is normal, despite the underlying distribution being exponential.

Simulations

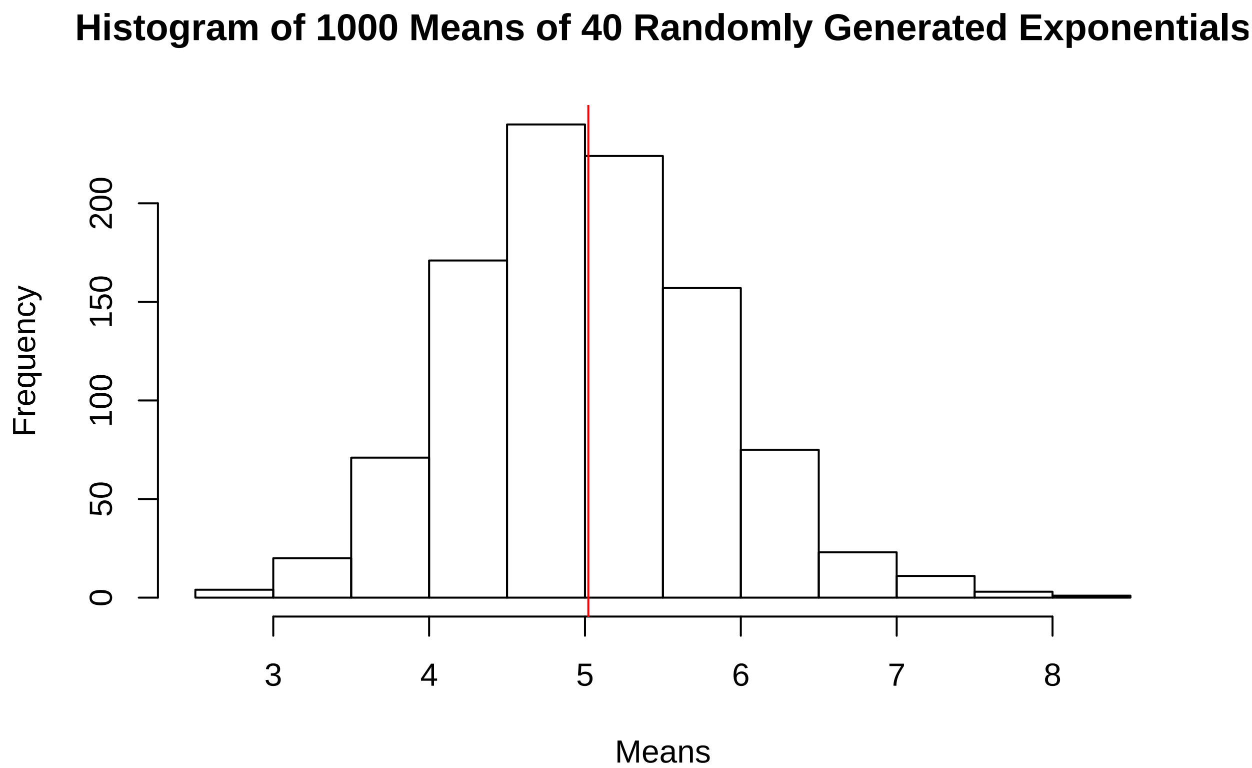
We simulated 1) an exponential distribution with n = 1000 and lambda = 0.2 to view the shape of the exponential distribution, and 2) 1000 exponential distributions with n = 40 and lambda = 0.2 to view the shape of the distribution of means of exponential distributions.

```
# Exponential distribution with n=1000
exp1000 <- rexp(1000, 0.2)

# 1000-large distribution of means of exponential distributions with n=40
mns = NULL
for (i in 1:1000) mns = c(mns, mean(rexp(40, 0.2)))
```

Sample Mean vs. Theoretical Mean

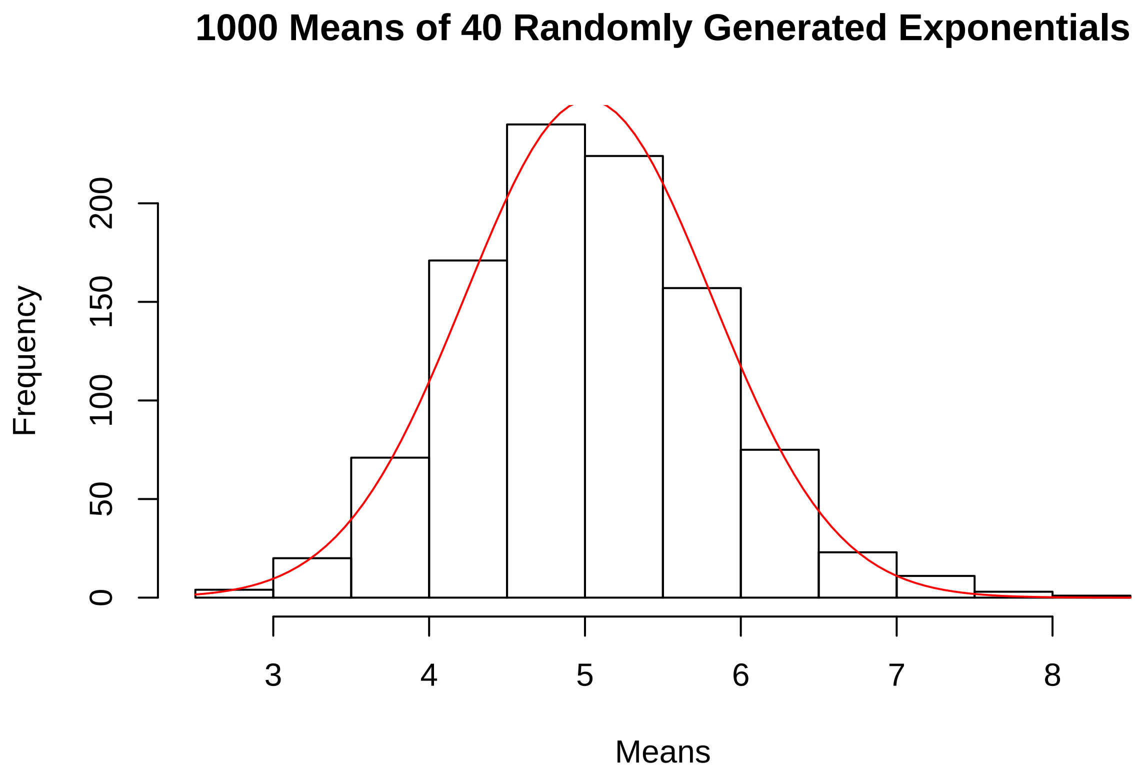
```
sample_mean <- mean(mns)
hist(mns, xlab='Means', main = 'Histogram of 1000 Means of 40 Randomly Generated Exponentials')
abline(v=sample_mean, col='red')
```



The theoretical mean of the distribution of the mean of 40 exponentials is 1/lambda = 5. On the plot above, the sample mean (5.0215588) is shown in red – it is very close to the theoretical mean.

Sample Variance vs. Theoretical Variance

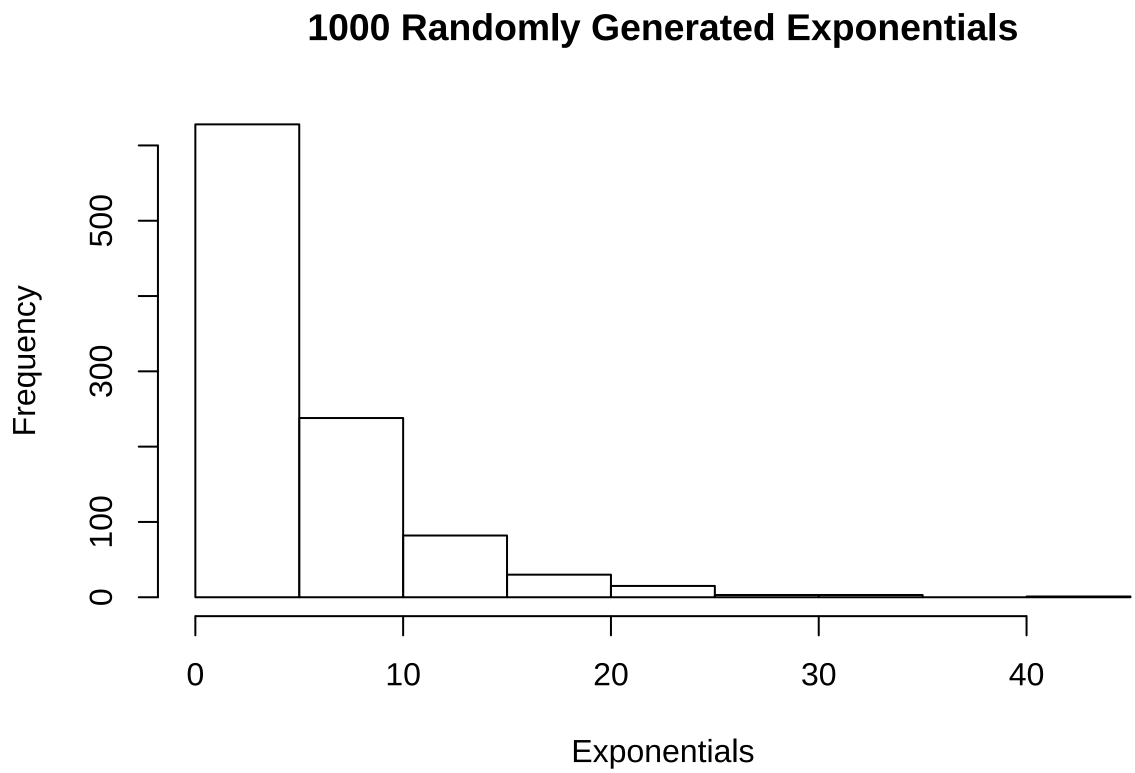
```
sample_variance <- var(mns)
theoretical_variance <- (5/sqrt(40))^2
hist(mns, xlab='Means', main = '1000 Means of 40 Randomly Generated Exponentials')
x <- seq(min(mns), max(mns), length=1000)
curve(dnorm(x, mean=sample_mean, sd=sqrt(theoretical_variance))*500, add=TRUE, col='red')
```



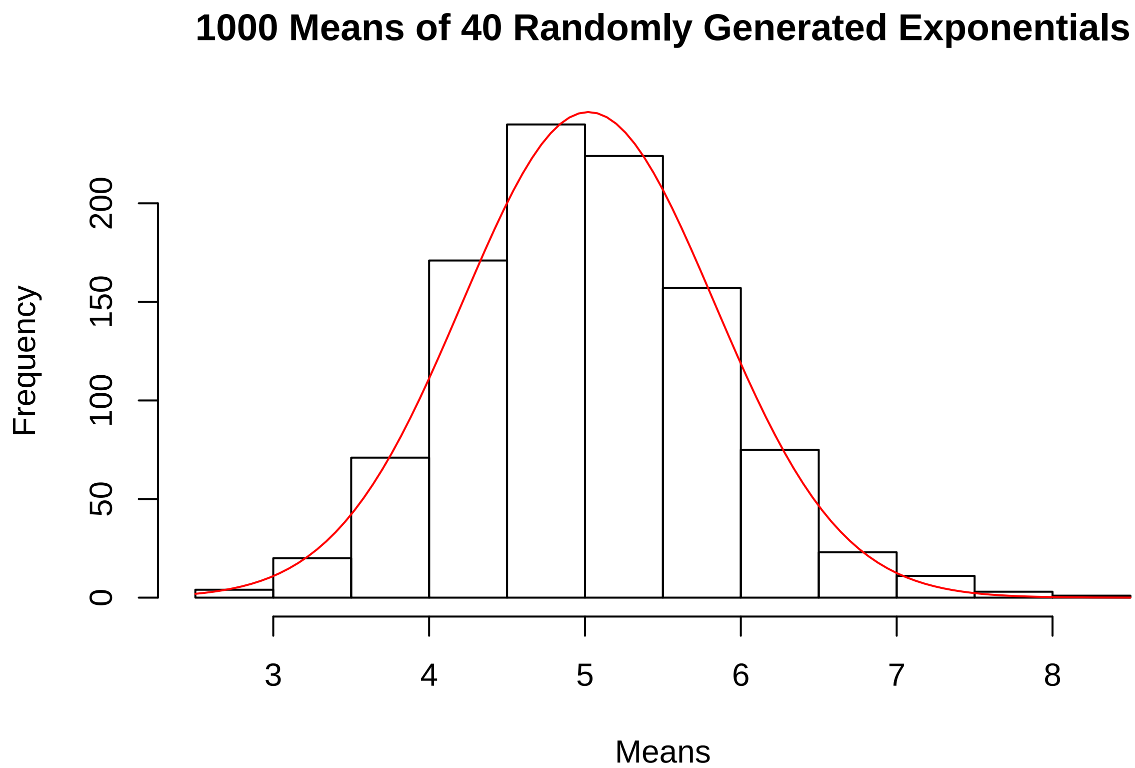
The theoretical variance of the distribution of means can be calculated as the square of the standard error of the underlying exponential distribution – in this case, the standard deviation of an exponential is 1/lambda = 5 and we used n = 40, so the theoretical variance is (5/sqrt(40))^2 = 0.625. Our sample variance (from 1000 simulations) is 0.6559676, which is very close. The plot above shows a histogram of our simulations with a curve showing a normal distribution centered at the same spot with the theoretical variance; here, we can graphically see that the variance of our simulated distribution is very close to the theoretical variance.

Distributions

```
hist(exp1000, xlab='Exponentials', main = '1000 Randomly Generated Exponentials')
```




```
hist(mns, xlab='Means', main = '1000 Means of 40 Randomly Generated Exponentials')
x <- seq(min(mns), max(mns), length=1000)
curve(dnorm(x, mean=sample_mean, sd=sqrt(sample_variance))*500, add=TRUE, col='red')
```



The figures above show the distribution of 1000 exponentials compared to the distribution of 1000 means of exponentials, in which we can clearly see that the two distributions have different shapes. The first is shaped as an exponential (with a long upper tail), while the second is shaped as a normal distribution. I have overlaid a normal distribution with mean = sample mean and variance = sample variance, to show that our distribution is approximately normal.