Statistical Inference Course Project: Simulation Excercise

By: Alexandra Hearn

Simulation To Verify Central Limit Theorem On Exponential Distribution

This project investigates the exponential distrbution and uses it to verify the central limit theorem. A thousand simulations are run each with 40 samples taken from an exponential distribution with rate parameter lambda = 0.2. Theoretical means and variances of the sample means is compared with the observed values and normality of the distribution is checked.

Simulations

40 random samples were drawn from an exponential distribution with rate parameter 0.2 using the rexp(n,lambda) function. The mean of each sample was taken and a vector of samples means was constructed for further analysis.

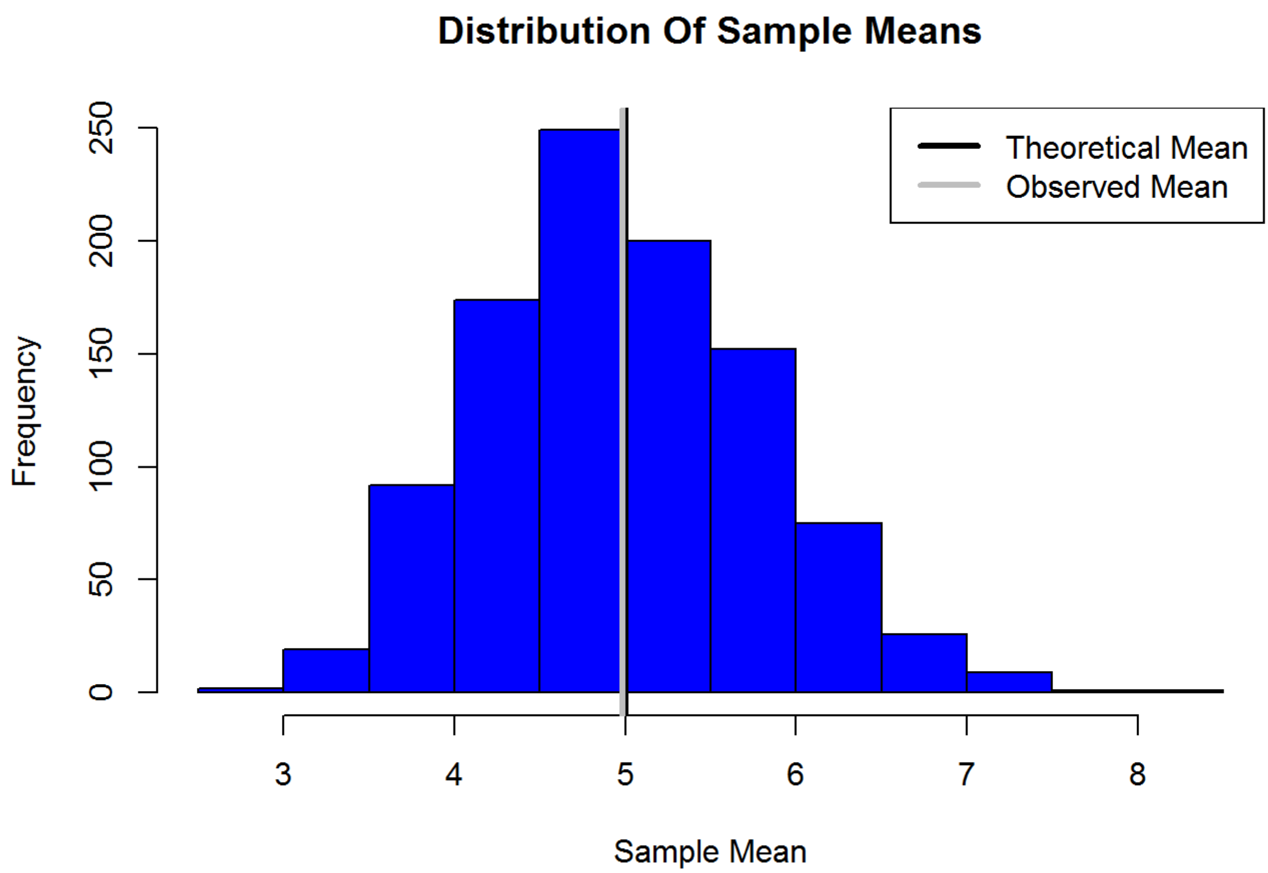
```
set.seed(1000)
noOfSimulations = 1000
n = 40
lambda = 0.2

means = NULL
for(i in 1:noOfSimulations)
{
  sample = rexp(n,lambda)
  means = c(means,mean(sample))
}
```

Comparison Of Mean Of Sample Mean To Theoretical Mean

According to CLT, the mean of sample means is normally distributed with mean equal to the actual theoretical mean of the original distribution. Thus, mean of samples means is expected to be equal to the theoretical mean of 1/lambda.

```
# Compare Mean Of Sample Means To Theoretical Mean
hist(means,main='Distribution Of Sample Means',xlab='Sample Mean',col='blue')
abline(v=1/lambda,col='black',lwd=3)
abline(v=mean(means),col='grey',lwd=3)
legend(x='topright',col=c('black','grey'),legend=c("Theoretical Mean","Observed Mean"),lwd=3)
```



Comparison

```
# Theoretical Mean
m1 = 1/lambda
print(m1)
```

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

```
# Observed Mean
m2 = mean(means)
print(m2)
```

```
## [1] 4.986963
```

```
# Absolute Difference
adiff = abs(m1 - m2)
print(adiff)
```

```
## [1] 0.01303661
```

We can see that the theoretical and observed values are very close to each other.

Comparison Of Variance Of Sample Mean To Theoretical Variance

According to CLT, the mean of sample means is normally distributed with standard deviation equal to the actual theoretical standard deviation of the original distribution divided by the square root of the sample size. Thus, variance of samples means is expected to be equal to the theoretical variance of 1/(lambda x lambda x 40).

```
# Compare Variance Of Sample Means To Theoretical Variance

# Theoretical Variance
v1 = (1/(lambda*lambda*n))
print(v1)
```

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

```
# Observed Variance
v2 = var(means)
print(v2)
```

```
## [1] 0.654343
```

```
# Absolute Difference
adiff = abs(v1-v2)
print(adiff)
```

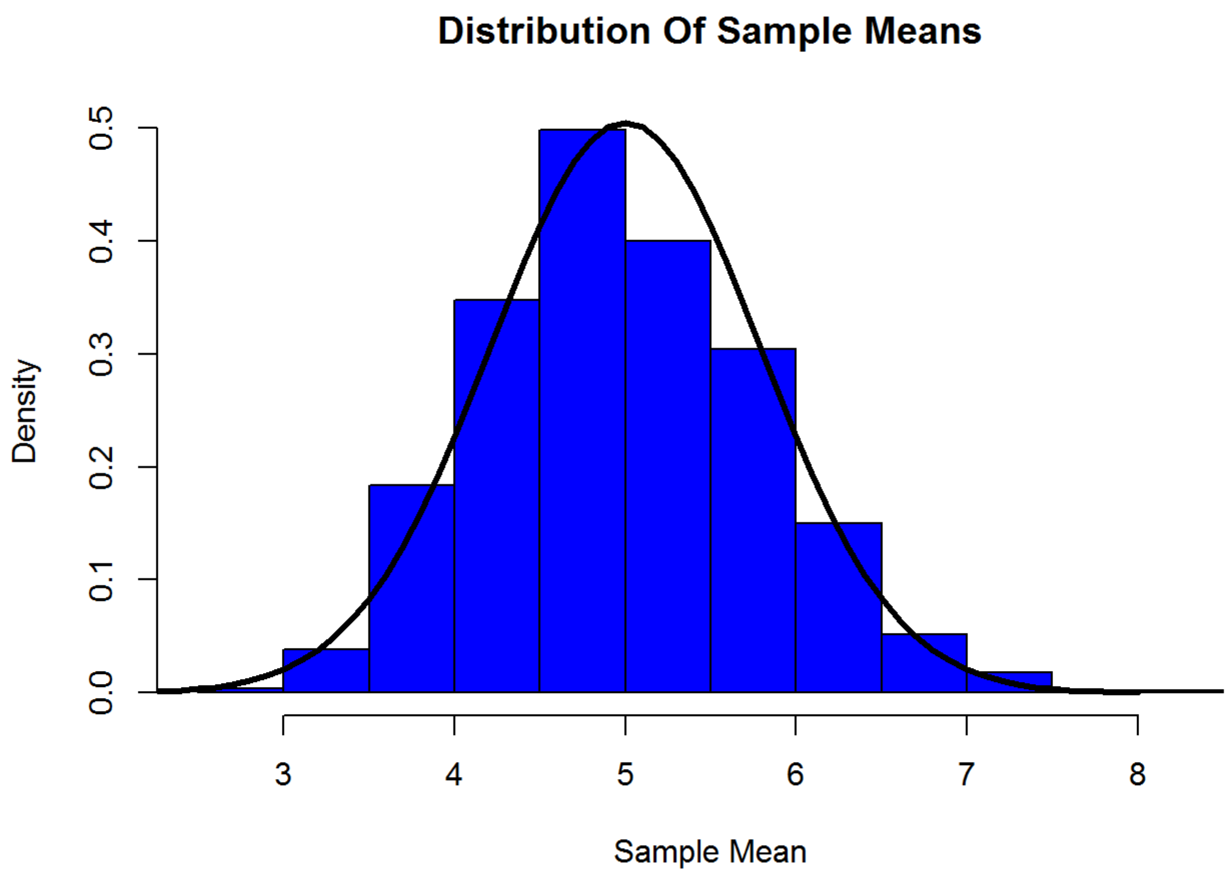
```
## [1] 0.02934304
```

We can see that the theoretical and observed values are very close to each other.

Approximate Normality Verification

We will plot a histogram of the sample means and overlay that with the theoretical normal distribution that is with mean (1/lambda) and variance 1/(lamda x lambda x 40).

```
# Showing Approximate Normality
hist(means,main='Distribution Of Sample Means',xlab='Sample Mean',col='blue',prob=T)
x = seq(-3,8,0.1)
y = dnorm(x,mean=m1,sd=sqrt(v1))
lines(x,y,lwd=3)
```



We can see from the plot that the distribution of sample means is approximately normal with the theoretical values of mean and standard deviation.

Conclusion

Based on the above results, we can see that the theoretical mean and variance of the sample means matches the observed values and also that the distribution of the observed sample means is approximately normal.