

Applying CLT on an Exponential distribution

Manish Mittal

Overview:

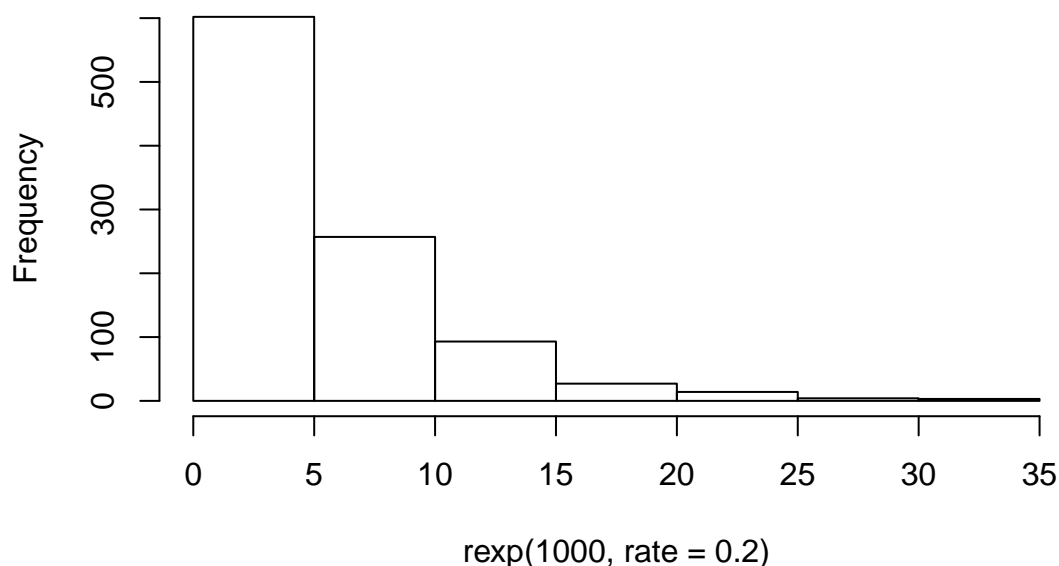
In this report I will apply the Central Limit Theorem (CLT) to Exponential distribution and show via a simulation that mean of 40 exponential distributions is normal as predicted by CLT. Further I will estimate the mean and variance expected from applying the CLT to the Exponential distribution.

Simulations:

First I computed the mean of the 40 samples from an exponential distribution with known rate. This computation was repeated 1000 times by using the “sapply” function and the data stored in variable sim. The plot below shows what a 1000 sample histogram from the exponential distribution will look like.

```
set.seed(1)
hist(rexp(1000,rate=0.2),main="Simulation of a histogram of exponential distribution")
```

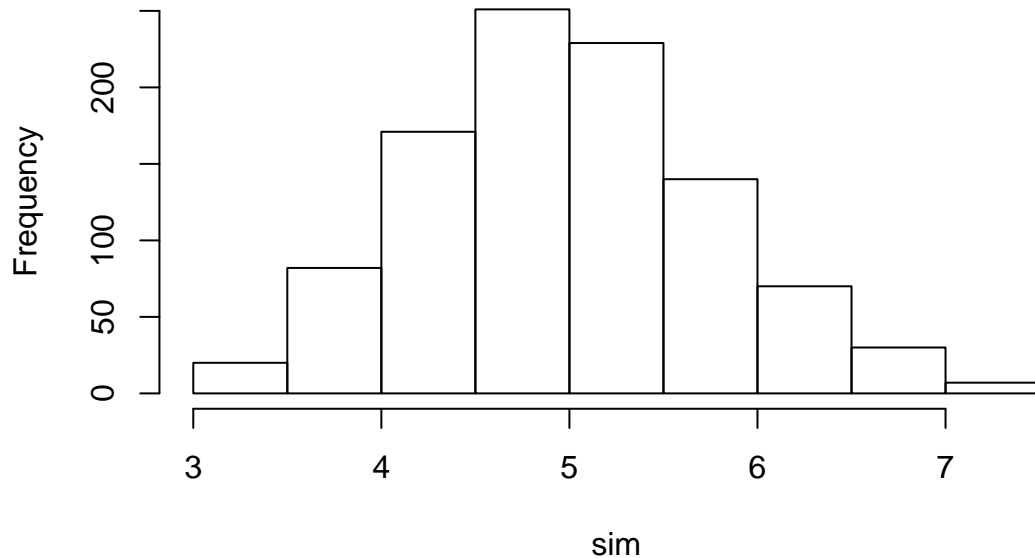
Simulation of a histogram of exponential distribution



The next plot below shows the result of 1000 simulations of average of 40 exponential samples with same rate as above plot. Note how this below plot looks fairly normal as opposed to above plot. This illustrates that via CLT that average of a large enough sample will follow a normal distribution. Next, we will estimate the mean and variance from the simulation and compare it with theoretical values.

```
sim <- sapply(1:1000,function(x) mean(rexp(40,rate=0.2)))
hist(sim, main="Simulation of a histogram of means of 40 exponentials")
```

Simulation of a histogram of means of 40 exponentials



The calculations below compute the mean and variance from the simulation and also from theoretical calculations.

```
theo_mean = 1/0.2
theo_var = (1/0.2)^2/40
paste("Simulation mean = ",mean(sim))
```

```
## [1] "Simulation mean = 4.98888195546751"
```

```
paste("Simulation variance = ",var(sim))
```

```
## [1] "Simulation variance = 0.611906551793589"
```

```
paste("Theoretical mean = ", theo_mean)
```

```
## [1] "Theoretical mean = 5"
```

```
paste("Theoretical variance = ",theo_var)
```

```
## [1] "Theoretical variance = 0.625"
```

The calculations below show what would be the 95% CI expected for an average of 40 samples which come from a population which is modeled by an exponential distribution. The limits are overlayed by blue lines on top of the simulated data and theoretical mean is shown in red.

```
paste0("95% CI for the expected mean of 40 exp distributions = (",theo_mean-1.96*sqrt(theo_var),",",theo
```

```
## [1] "95% CI for the expected mean of 40 exp distributions = (3.45048394651749,6.54951605348251)"
```

```
hist(sim, main="Simulation of a histogram of means of 40 exponentials")  
abline(v=theo_mean,col="red")  
abline(v=theo_mean-1.96*sqrt(theo_var),col="blue")  
abline(v=theo_mean+1.96*sqrt(theo_var),col="blue")
```

