Sample Means of Exponential Distributions Converge to Normal Distribution

Charin Polpanumas

February 5, 2559 BE

Overview

According to Central Limit Theorem (CLT), the sample means of any underlying distributions will be normally distributed given a sufficient number of iterations. This report demonstrates that it is also the case for exponential distribution, by comparing the sample mean and variance of the sample means to the theoretical mean and variance.

Simulations

We simulate the sample means of 40 exponential variables (lambda=0.2) for 1,000 iterations.

```r
set.seed(1111)
mns <- c()
n <- 1000
lambda <- 0.2
for (i in 1 : n) mns = c(mns, mean(rexp(40,lambda)))
```

Sample Mean/SD versus Theoretical Mean/SD

The sample mean of sample means converges to the theoretical mean 1/lambda. The sample standard deviation of sample means converges to the theoretical standard deviation (1/lambda)/sqrt(N) (N=40). The plot below shows where the sample mean/SD and the theoretical mean/SD are.
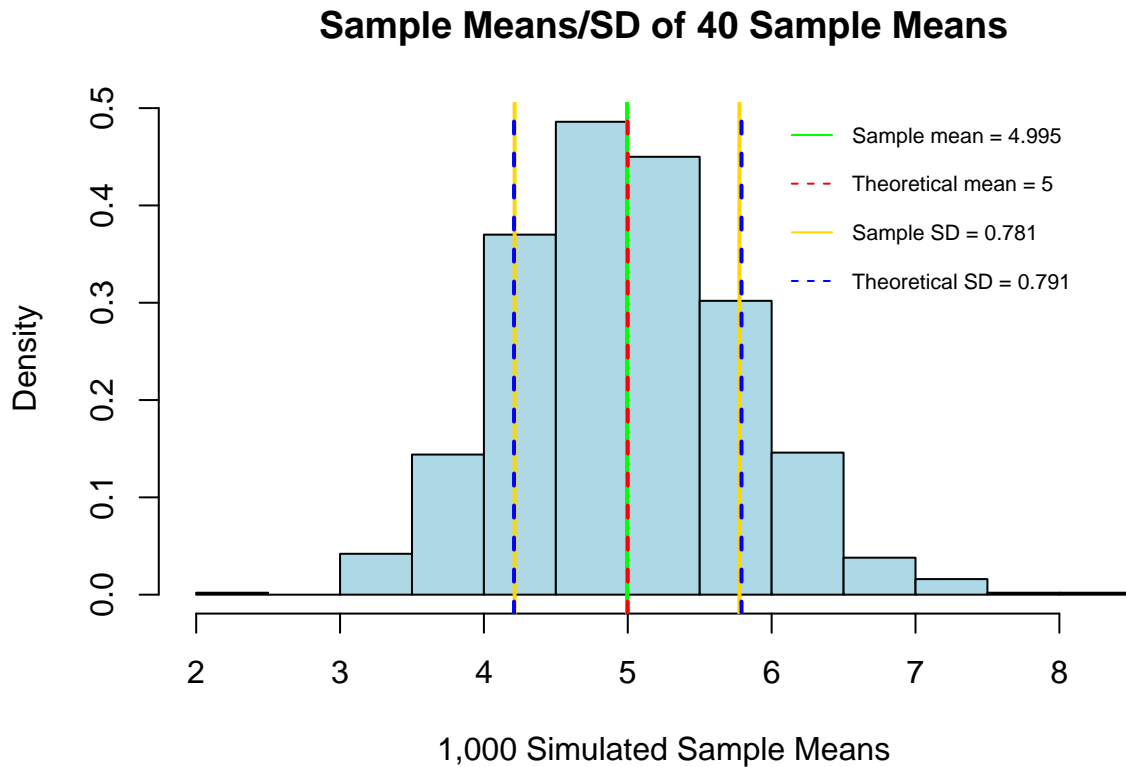
```r
#Mean
m <- mean(mns)
tm <- 1/lambda
hist(mns, col='lightblue', xlab='1,000 Simulated Sample Means', prob=1,
     main='Sample Means/SD of 40 Sample Means')
abline(v=m, col='green',lwd=2, lty=1)
abline(v=tm, col='red',lwd=2, lty=2)
legend(6,0.5,legend=paste('Sample mean =',round(m,3),sep=' '),cex=0.7,
       col = 'green', bty='n', lty=1)
legend(6,0.45,legend=paste('Theoretical mean =',tm,sep=' '),cex=0.7,
       col = 'red', bty='n', lty=2)

#SD
N=40
s <- sd(mns)
ts <- (1/lambda)/sqrt(N)
abline(v=m+c(-1,1)*s, col='gold',lwd=2, lty=1)
abline(v=tm+c(-1,1)*ts, col='blue',lwd=2, lty=2)
```

```r
legend(6,0.40,legend=paste('Sample SD =',round(s,3),sep=' '),cex=0.7,
       col = 'gold', bty='n', lty=1)
legend(6,0.35,legend=paste('Theoretical SD =',round(ts,3),sep=' '),
       cex=0.7, col = 'blue', bty='n', lty=2)
```
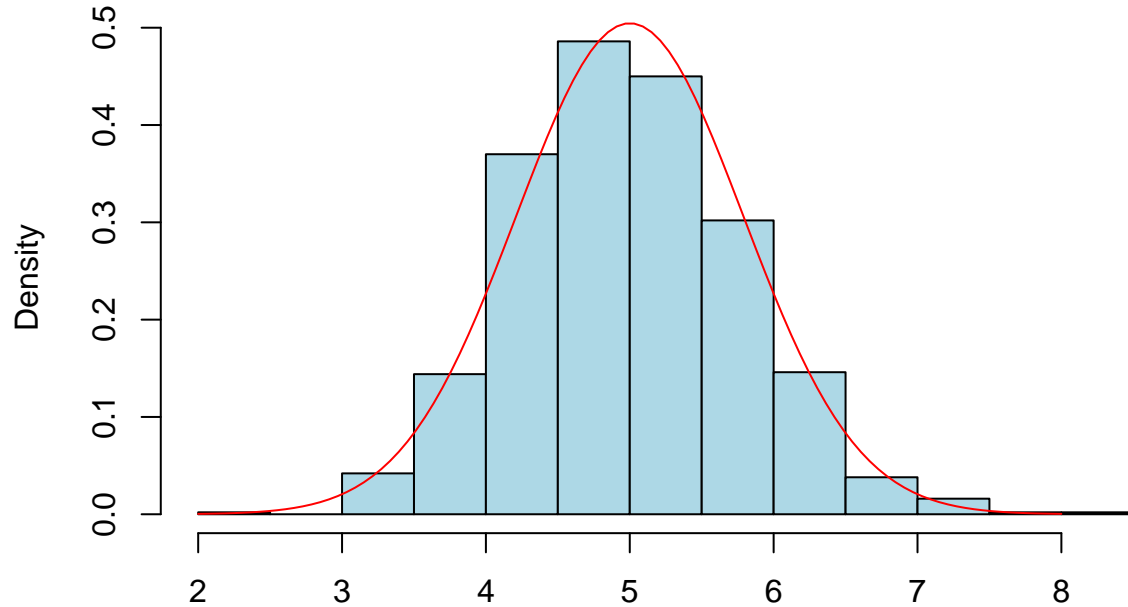

Distribution

The distribution of sample means converges to normal distribution (mean = 1/lambda, sd = 1/lambda/sqrt(N)).

```r
#Plot histogram of simulations
hist(mns, col='lightblue', xlab='1,000 Simulated Sample Means', prob=1,
     main='Sample Means of 40 Exponentially Distributed Variables')
#Overlay theoretical values
x<-seq(2,8, length=100)
hx<-dnorm(x,mean=tm, sd=ts)
lines(x,hx,type='l',col='red')
```

Sample Means of 40 Exponentially Distributed Variables

1,000 Simulated Sample Means

Q-Q plot demonstrates the normality of the distribution of sample means.

```r
qqnorm(mns, main = "Normal Q-Q Plot",
       xlab = "Theoretical Quantiles", ylab = "Sample Quantiles")
qqline(mns, col='red')
```

Normal Q−Q Plot