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.

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

Sample Mean versus Theoretical Mean

The sample mean of sample means converges to the theoretical mean 1/lambda. The sample mean of the sample means is:

```
m <- mean(mns)
m
```

```
## [1] 4.994883
```

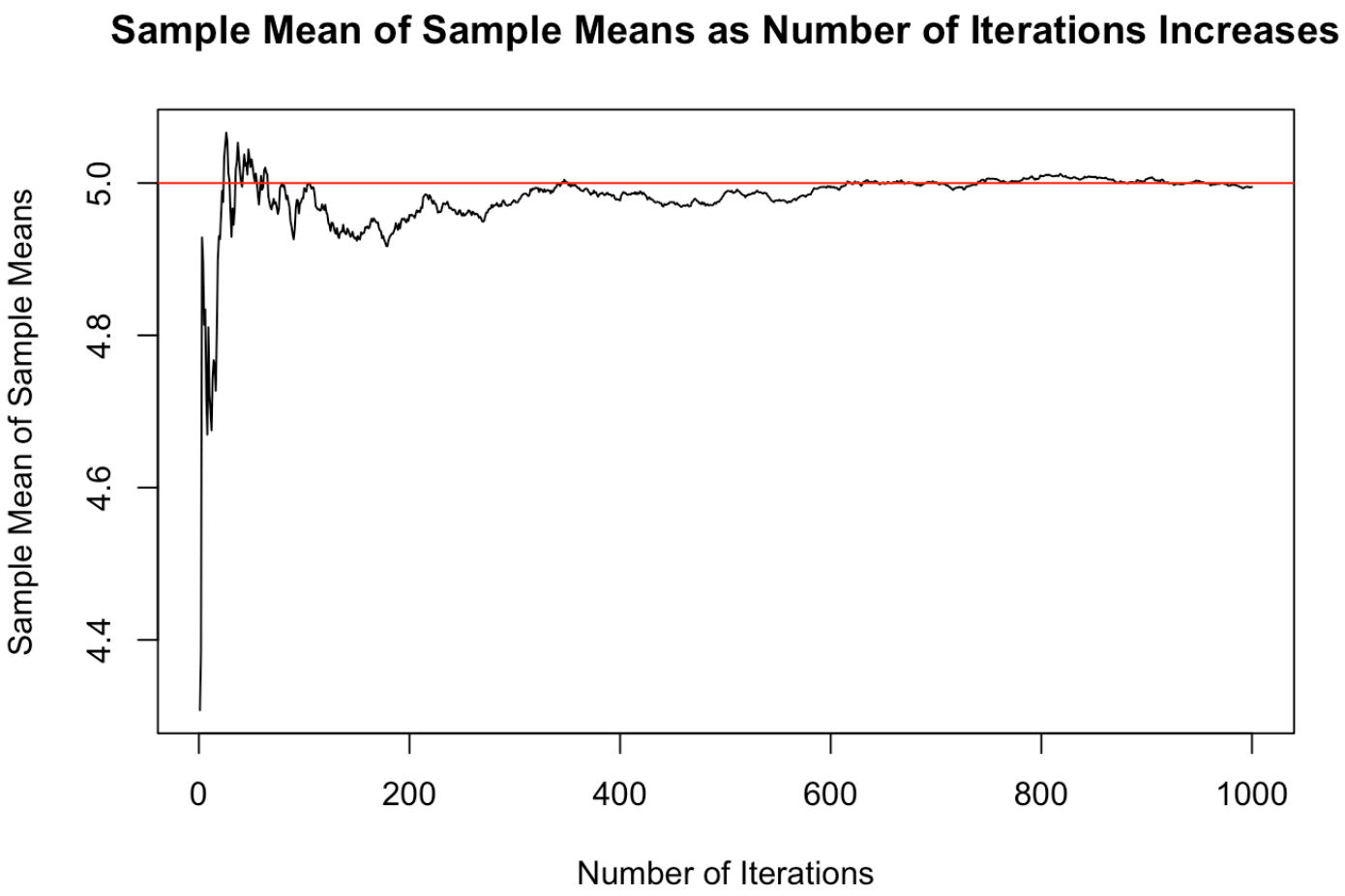
The theoretical mean is:

```
tm <- 1/lambda
tm
```

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

The more iterations, the more the sample mean approaches the theoretical mean.

```
plot(cumsum(mns)/1:n, type='l', main='Sample Mean of Sample Means as Number of Ite
rations Increases', xlab='Number of Iterations', ylab='Sample Mean of Sample Means
')
abline(h=tm, col='red')
```



Sample Variance versus Theoretical Variance The sample variance of sample means converges to the theoretical variance (1/lambda)^2/N (N=40). The sample variance of the sample means is:

```
v <- var(mns)
v
```

```
## [1] 0.6105269
```

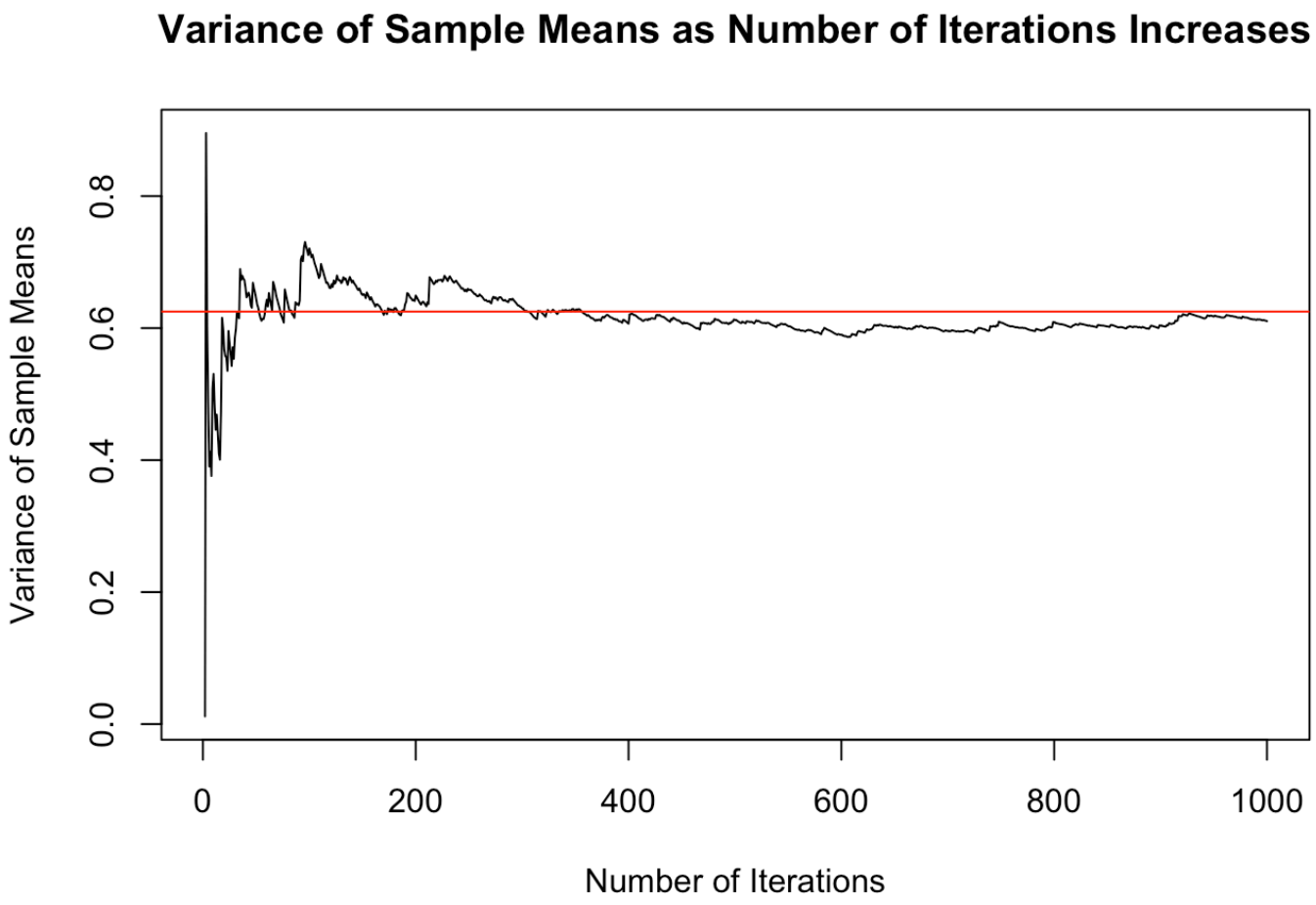
The theoretical variance is:

```
N <- 40
tv <- (1/lambda)^2/N
tv
```

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

The more iterations, the more the sample variance approaches the theoretical variance.

```
cv<-c()
for (i in 1:n){
    cv<-c(cv,var(mns[1:i]))
}
plot(cv, type='l', main='Variance of Sample Means as Number of Iterations Increase
s', xlab='Number of Iterations', ylab='Variance of Sample Means')
abline(h=tv, col='red')
```



Distribution

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

```
#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=sqrt(tv))
lines(x,hx,type='l',col='red')
```