Brief investigation on sampling from exponential distribution

Ilya Semenov

March 29, 2016

Overview

This is a brief illustration of Central Limit Theorem in action. Means and variances of 1000 simulated samples of length 40 exponential averages were compared with analytically obtained formulas for exponential mean and variance. Used exponential distribution with Lambda parameter 0.2.

Simulations

According to theory mean of exponential distribution with LAMBDA == 0.2 equals

```
lambda <- 0.2
1/lambda
```

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

So according to Central Limit Theorem (CLT) we can expect that mean of 1000 40-sized samples will approximately be equal to 5. Let check it.

```
mns <- NULL
set.seed(1001)
samplesize = 40
for (i in 1:1000) {
    mns <- c(mns, mean(rexp(samplesize, lambda)))
}
mean(mns)
```

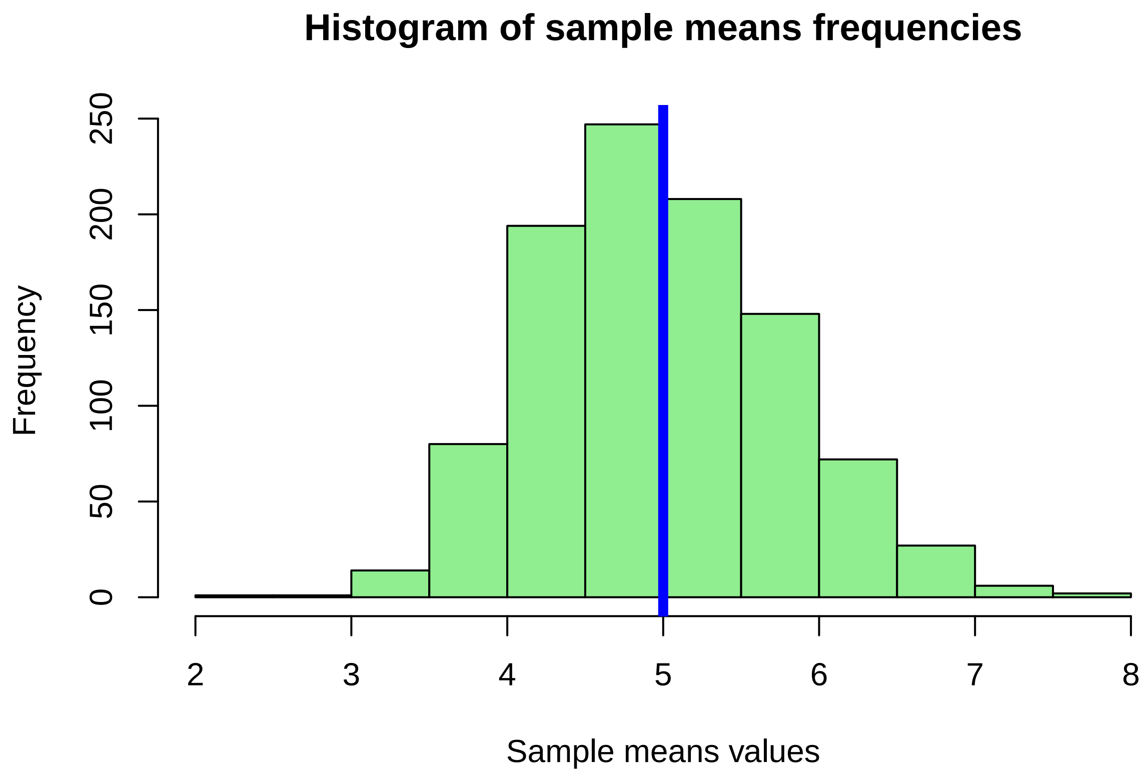
```
## [1] 4.973765
```

It proves our theoretical estimate is correct.

Sample mean vs theoretical mean

According to CLT sample means must be approximately normally distributed around theoretical mean of exponential distribution with lambda parameter 0.2, that equals 5. On a plot below blue vertical line represents theoretical mean.

```
hist(mns, main = 'Histogram of sample means frequencies',
     xlab = 'Sample means values', col = 'lightgreen')
abline(v=5, lwd=5, col='blue')
```



Sample variance vs theoretical variance

Theoretical standard deviation for exponential distribution samples of size 40 equals approximately 0.79.

```
thsd <- 1/lambda/sqrt(samplesize)
thsd
```

```
## [1] 0.7905694
```

Theoretical variance equals approximately 0.625.

```
thsd^2
```

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

Let check if simulated samples standard deviation and variance matche with those values.

```
sd(mns)
```

```
## [1] 0.7886699
```

```r
var(mns)
```

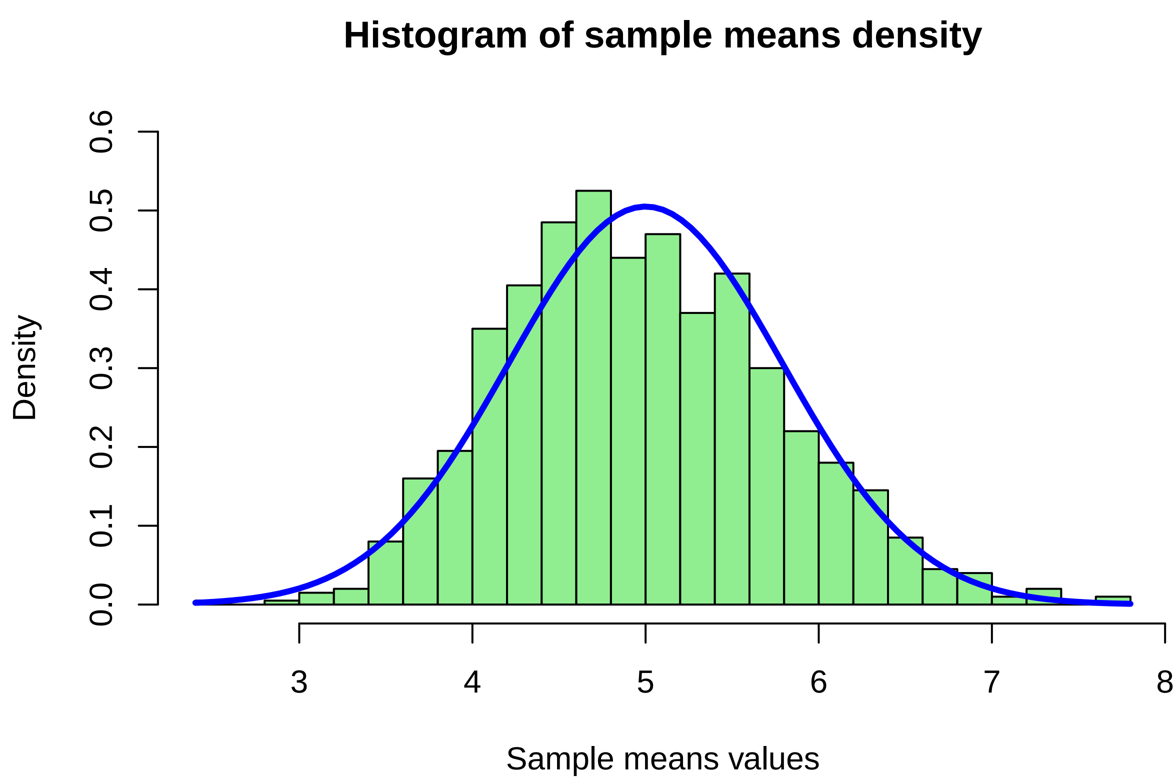
```
## [1] 0.6220001
```

So expected and obtained standard deviation and variance match.

Is distribution close to normal?

To answer this question we can graphically compare obtained samples mean distribution with normal distribution with parameters mean = 5, sd = 0.79

```r
hist(mns, breaks = 20, probability = TRUE, ylim = c(0, .6),
     main = 'Histogram of sample means density',
     xlab = 'Sample means values',
     col = 'lightgreen')
curve(dnorm(x, mean=5, sd = .79), lwd = 3, add = TRUE, col = 'blue')
```



So as seen from the figure sample means distribution fits with normal distribution perfectly.