Comparison of Expotential Distribution to Central Limit Theorem

Robert Jończy

25.07.2015

Overview

In this report I am going to compare expotential distribution to Central Limit Theorem (CLT). According to CLT, distribution of averages of iid variables (independent and identically distributed) becomes standard normal as sample size n increases. Meaning that $\bar{X}_n$ is approximately $N(\mu, \sigma^2/n)$

Due to report size limitations, I have turned off R code for plotting. Full repository for this report can be found here

Goal of this report is to show that:

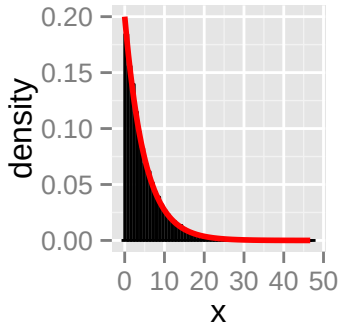
1. Sample mean compares to the theoretical mean of the distribution
2. Sample variance compares to the theoretical variance of the distribution
3. Distribution is approximately normal.

Simulations

Below, I draw the 40 samples for size $n = 1000$ from the exponential distribution with $\lambda = 0.2$.

```
set.seed(1234)
nosim <- 1000
lambda <- 0.2
n <- 40
data <- data.frame( x = rexp(nosim * n, lambda))
averages <- data.frame(x = apply(matrix(data$x,nosim),1,mean))
```

Below I plot histogram of simulated data and theoretical distribution.



Sample Mean vs. Theoretical Mean

Sample mean of 40 exponential distribution with 1,000 simulations:

```
(sample.mean <- mean(averages$x))
```

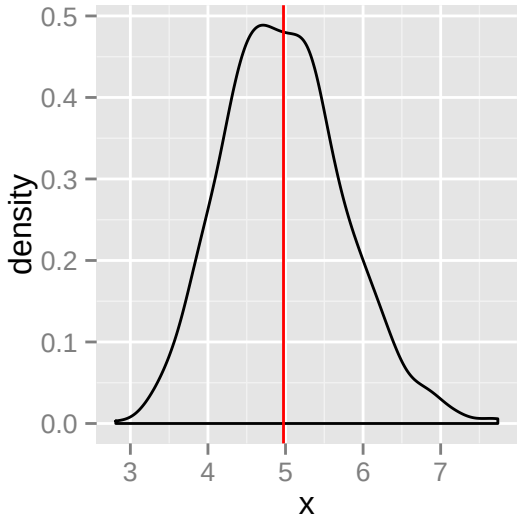
```
## [1] 4.974239
```

Mean from theoretical mean of the distribution

```
(theoretical.mean <- 1/lambda)
```

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

As we can see above simulation mean has converged to the theoretical value of 5. The sample mean of exponential distribution is very close to the theoretical center of the distribution (which is $1/\lambda$).



Sample Variance vs. Theoretical Variance

Simulated sample variance:

```
(sample.variance <- var(averages$x))
```

```
## [1] 0.5949702
```

Theoretical variance:

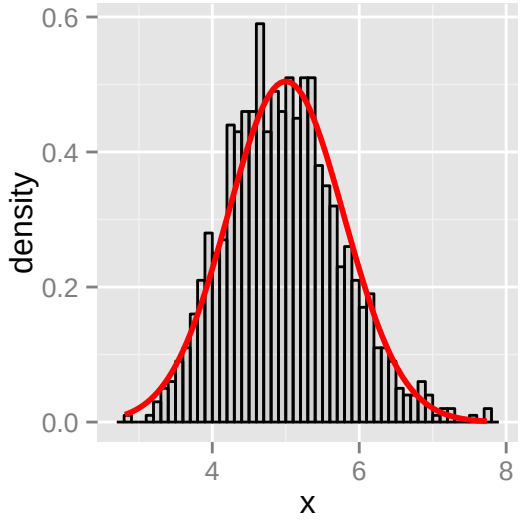
```
(theoretical.var <- 1 / (n * lambda^2))
```

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

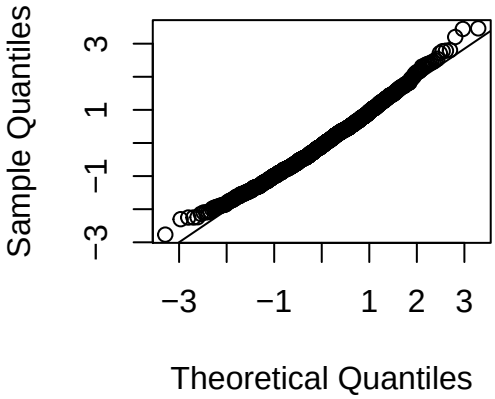
As we can see above the sample variance of exponential distribution and the theoretical variance are very close.

Comparison of expotential distribution to normal distribution

I am comparing distribution of averages with Normal Distribution ($N(\mu, \sigma^2/n)$) by ploting histogram of distribution and normal distribution.




```
#par(mar = c(5,5,0,0) + 0.5, mgp = c(2,1,0))
y <- (averages$x - mu) / (sigma / sqrt(40))  # normalize the data
qqnorm(y, main = NULL)
qqline(y)
```



On above QQ plot of the distribution of averages of 40 normalized exponential random variables we can easly see that the population is normally distributed.

Conclusion

Performed simulation shows that the distribution of the sample means is approximatly normal. A sample size of $n = 40$ seemed to be enough to guarantee that the distribution of sample means is normal. The mean of the expontential distribution of sample means is very close to the mean of the theoretical population. The same sample variance.