The Exponential Distribution and the Central Limit Theorem

Luis A Sandino

Created on 2016-03-17 23:20:05 using R version 3.2.3 (2015-12-10)

Synopsis

This project investigates the exponential distribution in R. Particularly, the distribution of averages of 40 exponentials is compared via simulation with the Central Limit Theorem (CLM).

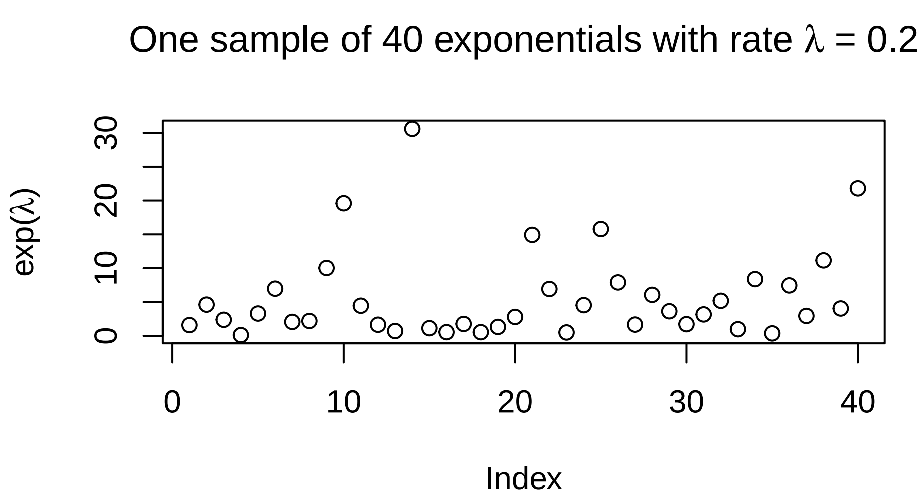
Simulations

Initialization:

```
set.seed(1984)  #Set the seed
lambda <- 0.2  #Set the exponential rate parameter
n <- 40  #Set the number of exponentials
sim <- 1000  #Set the number of simulations
```

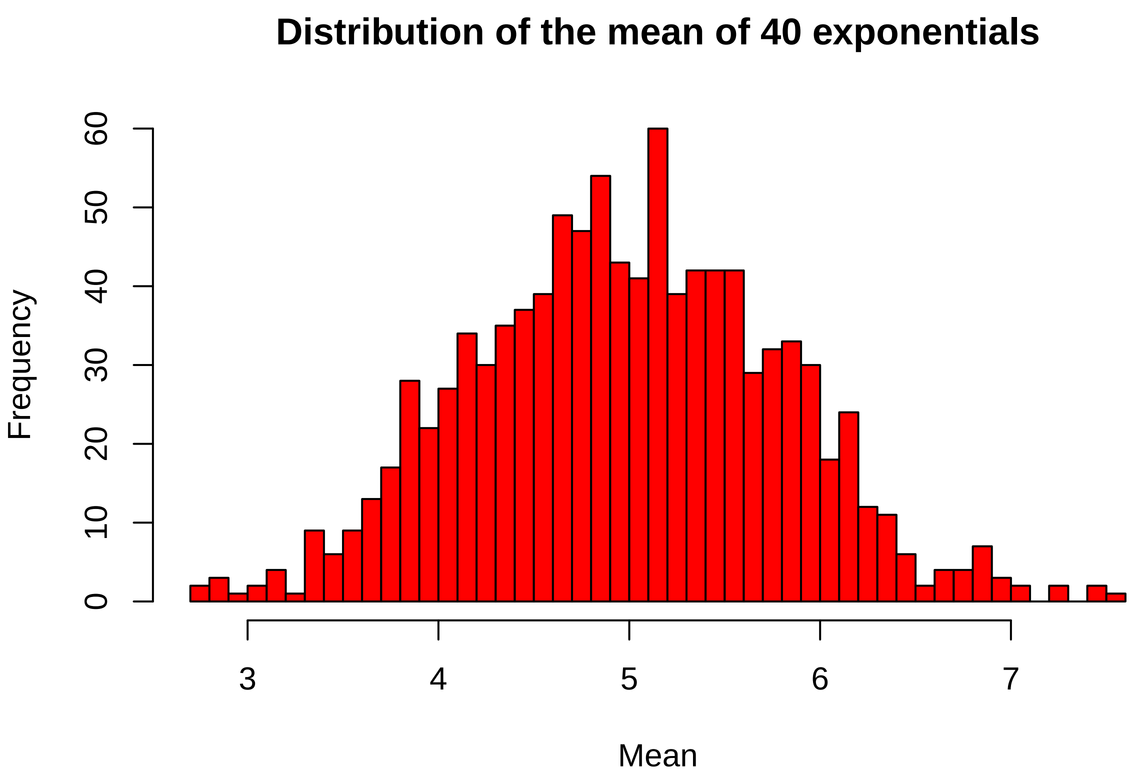
Results of one simulation of 40 exponentials:

```
plot(rexp(n, lambda), main = expression(paste("One sample of 40 exponentials with rate ",
    lambda, " = 0.2")), ylab = expression(paste("exp(", lambda, ")")))
```



Calculate the mean and variance of 40 exponentials in 1000 simulations and plot the distribution of the mean:

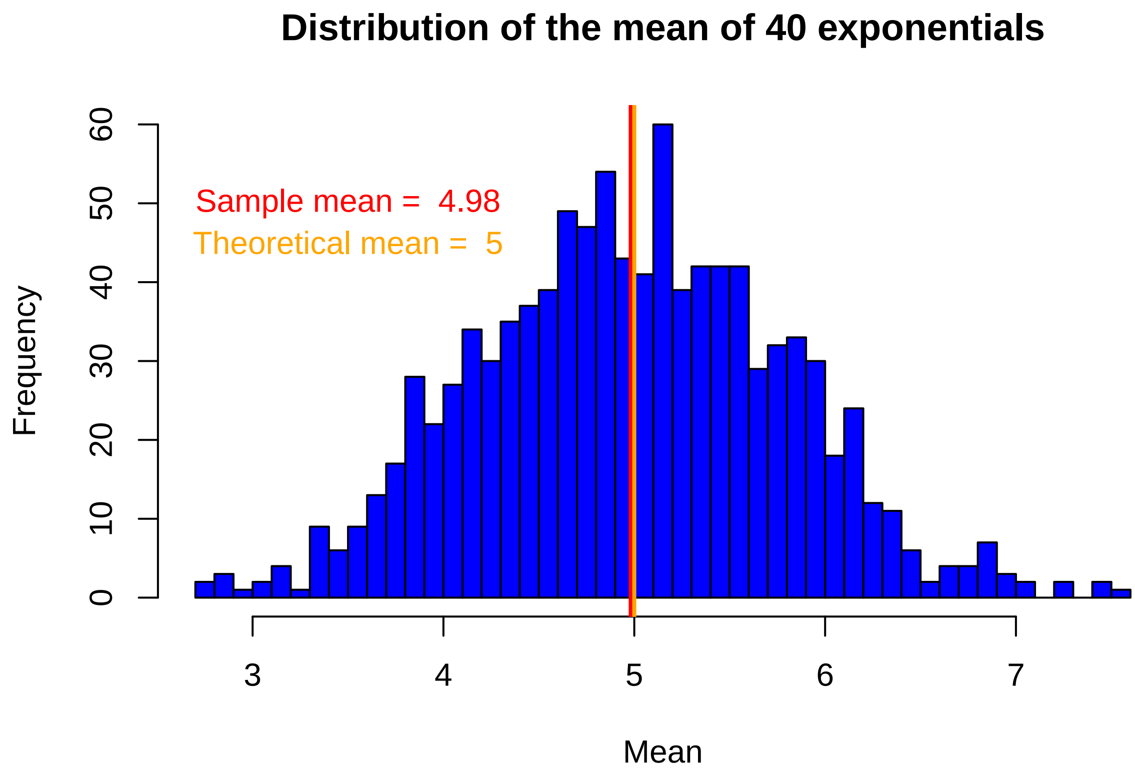
```
means <- NULL
vars <- NULL
for (i in 1:sim) {
    x <- rexp(n, lambda)
    means <- c(means, mean(x))
    vars <- c(vars, var(x))
}
hist(means, col = "red", breaks = 40, main = "Distribution of the mean of 40 exponentials",
    xlab = "Mean")
```



Sample Mean versus Theoretical Mean

In this section the average of sample mean is compared to the theoretical mean of the distribution. The mean of exponential distribution is $\frac{1}{\lambda}$

```
hist(means, col = "blue", breaks = 40, main = "Distribution of the mean of 40 exponentials",
    xlab = "Mean")
abline(v = mean(means), lwd = "2", col = "red")
abline(v = 1/lambda, lwd = "2", col = "orange")
text(3.5, 50, paste("Sample mean = ", round(mean(means), 2)), col = "red")
text(3.5, 45, paste("Theoretical mean = ", 1/lambda), col = "orange")
```



Sample Variance versus Theoretical Variance

In this section the average of sample variance is compared to the theoretical variance of the distribution. The variance of exponential distribution is $\left(\frac{1}{\lambda}\right)^2$

Calculate the average of sample variance

```
print(round(mean(vars), 2))
```

```
## [1] 24.64
```

Calculate the theoretical variance

```
print((1/lambda)^2)
```

```
## [1] 25
```

Distribution

The CLT states that, the mean of a sufficiently large number of iterates of independent random variables, each with a well-defined expected value and well-defined variance, will be approximately normally distributed, regardless of the underlying distribution.

In order to check this, plot the distribution of the mean together with the distribution of $N\left(\frac{1}{\lambda}, \frac{1}{\lambda \cdot \sqrt{n}}\right)$:

```r
hist(means, col = "blue", breaks = 40, main = "Distribution of the mean of 40 exponentials",
    xlab = "Mean", freq = FALSE)
x <- seq(-3, 3, length = 1000) * (1/lambda) + (1/lambda)
y <- dnorm(x, 1/lambda, (1/lambda)/sqrt(n))
lines(x, y, lwd = 4, col = "red")
text(3.5, 0.5, expression(paste("N", bgroup("(", paste(frac(1, lambda), ",",
    frac(1, lambda * sqrt(n))), ")"))), col = "red")
```