Peer-graded Assignment: Statistical Inference Course Project

Investigating and comparing the Central Limit Theorem and the exponential distribution in R

Koen Vermeulen

8-8-2019 [AANPASSEN]

## Overview

[In a few (2-3) sentences explain what is going to be reported on.]

## Simulations

[Include English explanations of the simulations you ran, with the accompanying R code. Your explanations should make clear what the R code accomplishes.]

Setting the seed, lambda, mean and standard deviation for the exponential distribution:

set.seed(2401)  
lambda <- 0.2  
mean <- 1/lambda  
sd <- 1/lambda  
exp <- 40  
n <- 1000

### Running the simulations

sampleMeans <- NULL  
for (i in 1:n) sampleMeans <- c(sampleMeans, mean(rexp(exp, lambda)))

### Sample Mean versus Theoretical Mean

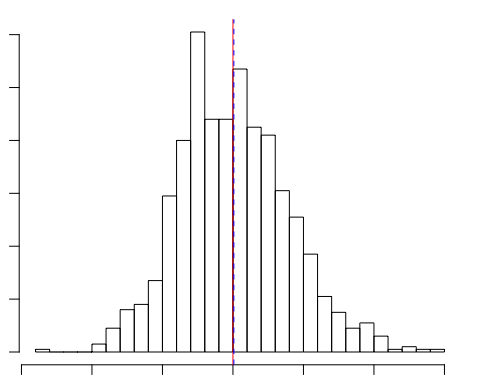
Calculating the sample mean:

meanS <- mean(sampleMeans)  
meanS

[1] 5.012893

The sample mean is 5.012893, the theoretical mean is 5. There is a small difference of 0.012893.

#### Histogram of the comparison



The observed sample mean is plotted as a dashed blue line and the theoretical mean is plotted as a red line.

### Sample Variance versus Theoretical Variance

Calculating the variance from the sample means:

varS <- var(sampleMeans)  
varS

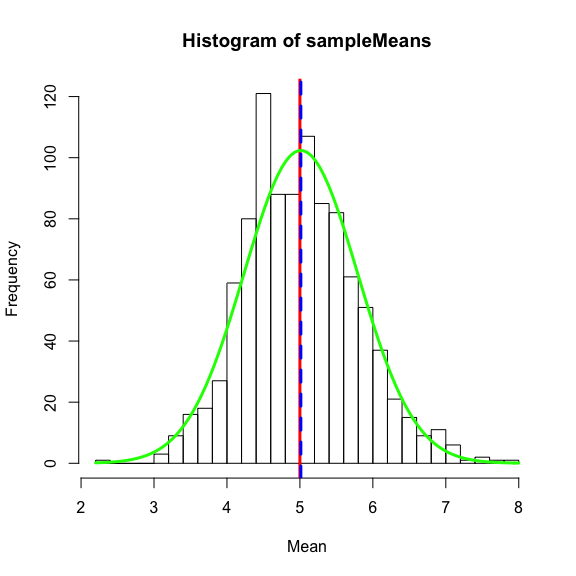
[1] 0.6073928

varT <- (lambda \* sqrt(exp))^-2

The variance of the sample means is 0.6073928, the theoretical mean of an exponential distribution is (lambda \* sqrt(n))^-2. In this case this is 0.625. Again, the difference is only small of -0.0176072.

### Distribution

#### Histogram of means



The histogram above is the same as the earlier one, expanded with the theoretical distribution. In the graph it is shown as the green line. As can be seen the distribution of the sample means is approximately normally distributed.