Statistical_Inference_1

Ajith Masthan

May 17, 2020

Statistical Inference Part-1

Overview

Explore the exponential distribution in R and compare it with the Central Limit Theorem. The exponential distribution is simulated in R with rexp(n, lambda) where lambda is the rate parameter. The mean and standard deviation of exponential distribution is 1/lambda. lambda = 0.2 is used for all of the simulations. Distribution of averages of 40 exponentials is investigated with a thousand simulations.

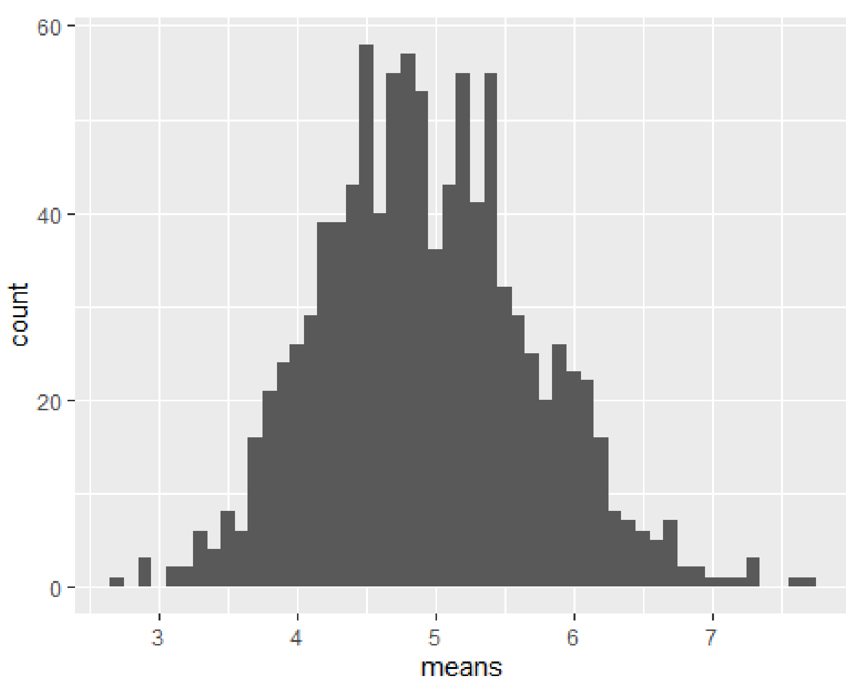
Simulations

```
# load neccesary libraries
library(ggplot2)

# set constants
lambda <- 0.2 # lambda for rexp
n <- 40 # number of exponetials
numberOfSimulations <- 1000 # number of tests

# set the seed to create reproducability
set.seed(2866)

# run the test resulting in n x numberOfSimulations matrix
exponentialDistributions <- matrix(data=rexp(n * numberOfSimulations, lambda),
nrow=numberOfSimulations)
exponentialDistributionMeans <- data.frame(means=apply(exponentialDistributions, 1,
mean))
```



Sample Mean versus Theoretical Mean

The expected mean mu of a exponential distribution of rate lambda is

mu = 1/lambda

```
mu <- 1/lambda
mu

## [1] 5
```

X_bar is the average sample mean of 1000 simulations of 40 randomly sampled exponential distributions.

```
meanOfMeans <- mean(exponentialDistributionMeans$means)
meanOfMeans

## [1] 4.941941
```

The expected mean is close to average sample mean

Sample Variance versus Theoretical Variance

The expected standard deviation sigma of a exponential distribution of rate lambda is

sigma = (1/lambda)/sqrt(n)

```
sd <- 1/lambda/sqrt(n)
sd

## [1] 0.7905694
```

The variance Var of standard deviation sigma is

Var=square(sigma)

```
Var <- sd^2
Var

## [1] 0.625
```

Let Var_x is the variance of the average sample mean of 1000 simulations of 40 randomly sampled exponential distribution with standard deviation sigma_x

```
sd_x <- sd(exponentialDistributionMeans$means)
sd_x

## [1] 0.7710459

Var_x <- var(exponentialDistributionMeans$means)
Var_x

## [1] 0.5945119
```

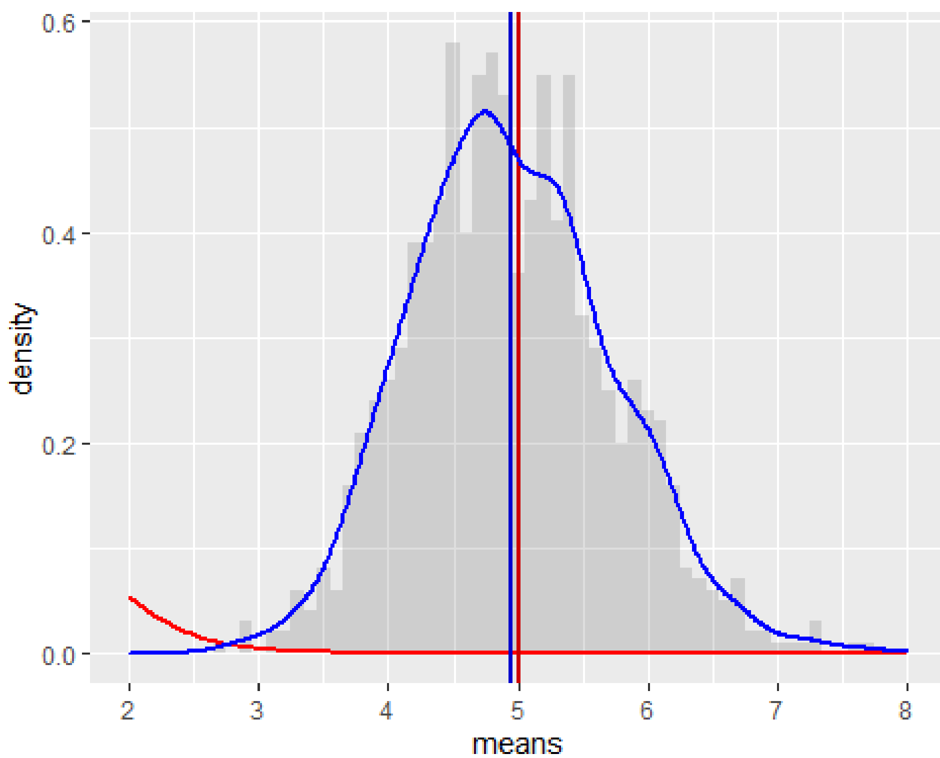
The Standard deviation and Variance are close

Distribution

Population Mean & Standard Deviation with a normal distribution of the expected values.

```
## Warning: Ignoring unknown parameters: arg

## Warning: Removed 2 rows containing missing values (geom_bar).
```



Observarion : The calculated distribution of means of random sampled exponantial distributions matches with the normal distribution with the expected values based on the given lamba