Course Proj 1

Leonardo78rs

05/04/2019

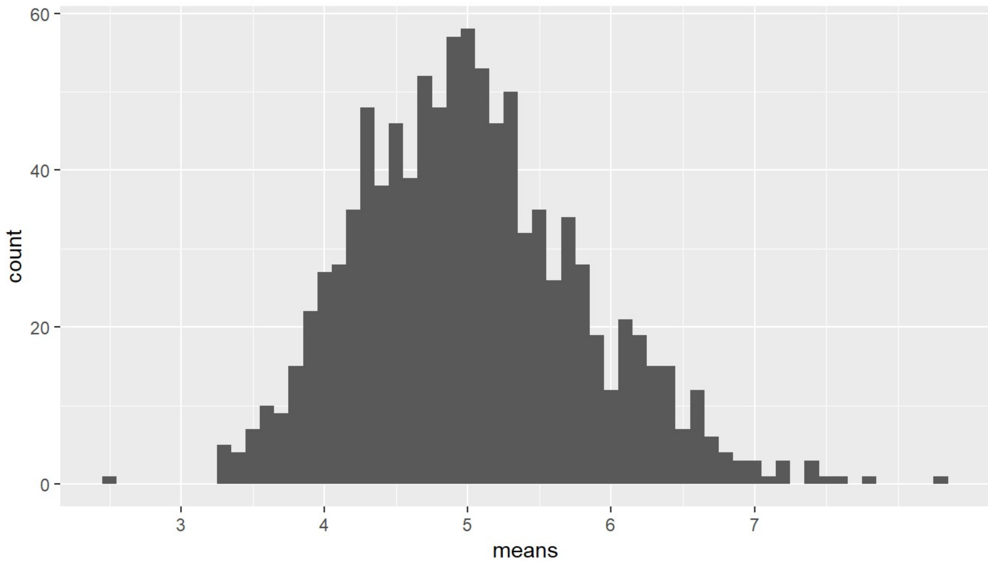
Statistical Inference Course Project 1

Overview

Part 1: Simulation Exercise Instructionsmenos In this project you will investigate the exponential distribution in R and compare it with the Central Limit Theorem. The exponential distribution can be simulated in R with rexp(n, lambda) where lambda is the rate parameter. The mean of exponential distribution is 1/lambda and the standard deviation is also 1/lambda. Set lambda = 0.2 for all of the simulations. You will investigate the distribution of averages of 40 exponentials. Note that you will need to do a thousand simulations.

Simulations (1000 random of 40 exp)

```r
# load ggplot2
library(ggplot2)
# set constants
lambda <- 0.2
n <- 40 # exponetials
QuantSimul <- 1000 # quant of tests
# set the seed to create reproducability
set.seed(1507446545)
# Running in N versus QuantSimul matrix
ExpDistrib      <- matrix(data=rexp(n * QuantSimul, lambda), nrow=QuantSimul)
ExpDistribMeans <- data.frame(means=apply(ExpDistrib, 1, mean))
```



Giving sample Mean versus Theoretical Mean

Note that the expected mean μ of a exponential distribution of rate λ is

$\mu = \frac{1}{\lambda}$

```r
mu <- 1/lambda
mu
```

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

Let $\bar{X}$ be the average sample mean of 1000 simulations of 40 randomly sampled distributions.

```
GrandMeans <- mean(ExpDistribMeans$means)
GrandMeans
```

```
## [1] 5.020541
```

The expected mean is very close of the avarage sample mean.

Sample Variance versus Theoretical Variance

The expected standard deviation σ of a exponential distribution of rate λ is

$\sigma = \frac{1/\lambda}{\sqrt{n}}$

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

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

The $Variance$ of standard deviation σ is

$Variance = \sigma^2$

```
Variance <- sd^2
Variance
```

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

- $Variance_x$ is the variance of the average sample mean of 1000 simulations of 40 randomly distributed;
- σ_x the corresponding standard deviation.

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

```
## [1] 0.7946587
```

```
Variance_x <- var(ExpDistribMeans$means)
Variance_x
```

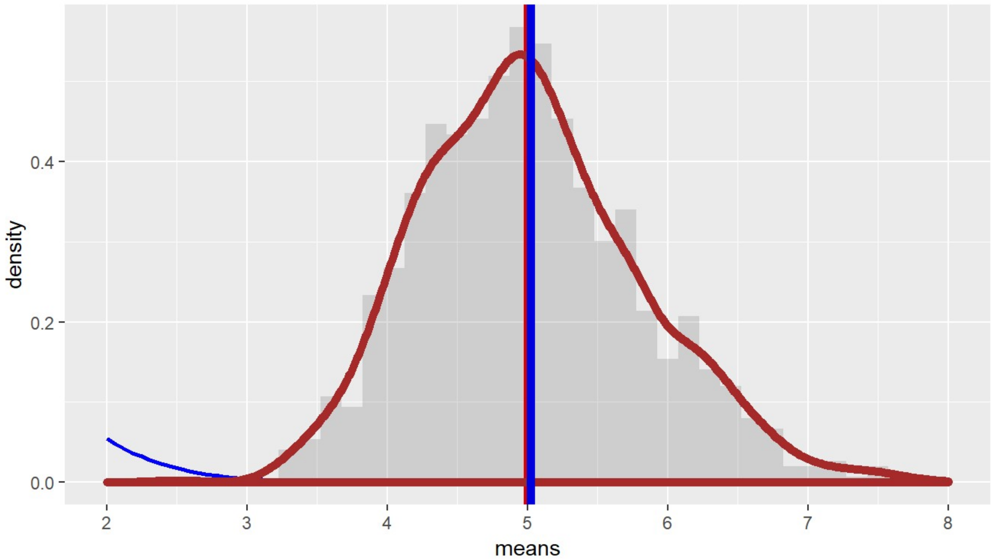
```
## [1] 0.6314825
```

The standard deviations are almost close together.

Distribution and Population means

The population means & standard deviation with a normal distribution of the expected values:

```
# plot
ggplot(data = ExpDistribMeans, aes(x = means)) +
  geom_histogram(binwidth=0.15, aes(y=..density..), alpha=0.2) +
  stat_function(fun = dnorm, arg = list(mean = mu , sd = sd), colour = "blue", size=1) +
  geom_vline(xintercept = mu, size=2, colour="#DD0000") +
  geom_density(colour="brown", size=2) +
  geom_vline(xintercept = GrandMeans, size=2, colour="#0000DD") +
  scale_x_continuous(breaks=seq(mu-3,mu+3,1), limits=c(mu-3,mu+3))
```



According to the graph, the calculated distribution of the means of the random exponential distributions of the sample overlaps the normal distribution with the expected values based on the lambda variable.