Project 1

Hana Kysela

May 21, 2016

Project 1 - Simulation excercise

Overview

In this project I will investigate the exponential distribution in R and compare it with the Central Limit Theorem.

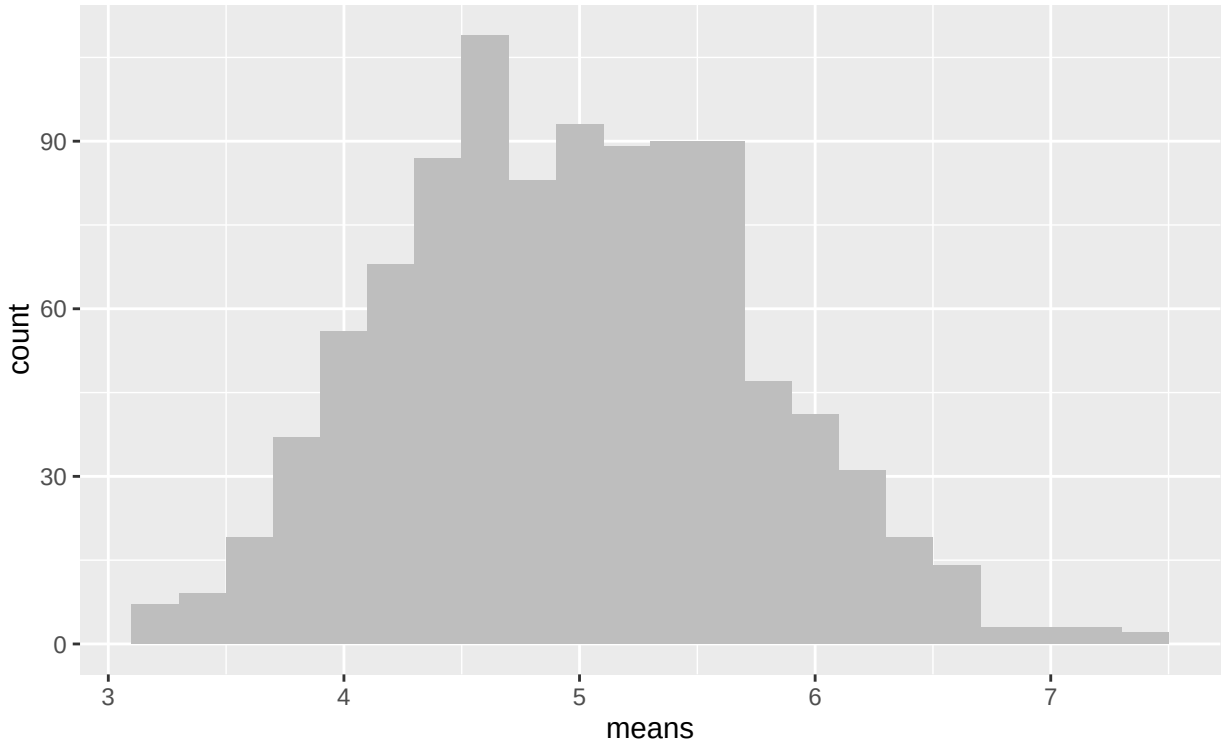
I am going to Show the sample mean and compare it to the theoretical mean of the distribution. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution. Show that the distribution is approximately normal.

Simulation

```
set.seed(1234) #allows reproducibility

# packages management
if (!require("pacman")) install.packages("pacman")
  pacman::p_load(ggplot2, knitr)

# The exponential distribution can be simulated in R with rexp(n, lambda) where lambda is the rate para
# 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

lambda <- 0.2   # lambda
n <- 40         # exponentials
sims <- 1000     # simulations

ExpMeans<-data.frame(sapply((1:sims), function(x) {mean(rexp(n, lambda))}))
colnames(ExpMeans)<- "means"
```

Histogram of the means

```
ggplot(data = ExpMeans, aes(x = means)) +
  geom_histogram(binwidth=0.2, fill = "grey") +
  scale_x_continuous(breaks=round(seq(min(ExpMeans$means), max(ExpMeans$means), by=1)))
```



1. Show the sample mean and compare it to the theoretical mean of the distribution.

Expected (theoretical) mean μ is $\frac{1}{\lambda}$

```
1/lambda
```

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

The actual mean from the simulation is

```
mean(ExpMeans$means)
```

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

Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.

Expected standard deviation of sample with standard deviation σ is $\sigma = \frac{1/\lambda}{\sqrt{n}}$ Expected (theoretical) variance Var of this sample is $Var = \sigma^2$

```
sd<-1/lambda/sqrt(n)
var<-sd^2
var
```

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

The actual variance from the simulation is

```
sd<-sd(ExpMeans$means)
var<-sd^2
var
```

```
## [1] 0.5706551
```

3. Show that the distribution is approximately normal.

The chart shows the histogram of simulated means and the theoretical values are in red, the simulation is in blue.

```
ggplot(data = ExpMeans, aes(x = means)) +
  geom_histogram(aes(y=..density..), binwidth = 0.2, fill = "grey") +
  stat_function(fun = dnorm, args = list(mean = 5 , sd = sd(ExpMeans$means)), colour = "red", size=1) +
  stat_function(fun = dnorm, args = list(mean = mean(ExpMeans$means) , sd = sd(ExpMeans$means)), colour
  geom_vline(xintercept = 5, size=1, colour="red") + # mean theory
  geom_vline(xintercept = mean(ExpMeans$means), size=1, colour="blue") + # mean sim

  scale_x_continuous(breaks=round(seq(min(ExpMeans$means), max(ExpMeans$means), by=1)))
```