Peer-graded Assignment: Statistical Inference Course ProjectPart 1

Alejandro Coy

2019-02-15

# Peer-graded Assignment: Statistical Inference Course Project

The project consists of two parts:

1. A simulation exercise.
2. Basic inferential data analysis.

## Part 1: Simulation Excercise :

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.

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

Theoretical Mean of distribution is 1/ lambda if lmabda is 0.2, the theoretical mean is 5

set.seed(150)  
lambda = 0.2  
n = 40 #number of exponentials  
n\_simulations =1000  
  
exp\_mean = NULL  
for(i in 1:n\_simulations){  
 exp\_mean = c(exp\_mean,mean(rexp(n,lambda)))  
}  
mean(exp\_mean)

## [1] 5.00215

Result indicate that the mean of 40 exponentials for 10000 simulations is close to the theoretical mean of 5.

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

The theoretical standart deviation is 1/lambda/sqrt(n) if, lamda is 0.2, the standart deviations is 5 and theoretical variance is 0.79

theo\_variance <- (1/lambda/sqrt(n))^2  
variance\_exp <- var(exp\_mean)  
variance\_exp

## [1] 0.6498806

theo\_variance

## [1] 0.625

In this case the simulation is a little further.

### Show that the distribution is approximately normal.

To probe normality the exp\_mean could be plotted in a histrogram

library(ggplot2)  
  
data <- data.frame(exp\_mean,n)  
ggplot(data = data,aes(exp\_mean))+  
 geom\_histogram(aes(y=..density..),fill = "steelBlue",binwidth = 0.2)+  
 stat\_function(fun = dnorm,args= list(mean=1/lambda,sd=1/lambda/sqrt(40)),color ="red")

