Final project (Statistical Inference)

Robinson Montes Gómez

11/15/2020

PART 1: SIMULATION EXCERCISE

Let's investigate the exponential distribution and compare it with the Central Limit Theorem. The first step is to create a random data with the rexp function with lambda 0.2

```
exp_data <- as.data.frame(replicate(40, mean(rexp(1000, 0.2))))
head(1/exp_data, 10)
```

```
##    replicate(40, mean(rexp(1000, 0.2)))
## 1                              0.2067405
## 2                              0.2055213
## 3                              0.1989486
## 4                              0.2074402
## 5                              0.2000963
## 6                              0.1994809
## 7                              0.2022394
## 8                              0.1942802
## 9                              0.1977445
## 10                             0.2002640
```

Calculating the means of the 40 simulation of size 1000, we can see that all means are approximately 0.2. Let's compute the average of the 40 means

```
mean_sample <- 1/mean(exp_data$`replicate(40, mean(rexp(1000, 0.2)))`)
round(mean_sample, 2)
```

```
## [1] 0.2
```

The value 0.2 is the average of the 40 means calculated in the simulations rounding to 2 decimals, it is the expected value of theoretical mean.

Now we are going to check the standard desviation of the simulations

```
exp_data_sd <- as.data.frame(replicate(40, sd(rexp(1000, 0.2))))
head(1/exp_data, 10)
```

```
##    replicate(40, mean(rexp(1000, 0.2)))
## 1                              0.2067405
## 2                              0.2055213
## 3                              0.1989486
## 4                              0.2074402
## 5                              0.2000963
## 6                              0.1994809
## 7                              0.2022394
## 8                              0.1942802
## 9                              0.1977445
```

```
## 10                                0.2002640
```

```
sd_sample <- 1/mean(exp_data_sd$`replicate(40, sd(rexp(1000, 0.2)))`)
round(mean_sample, 2)
```

```
## [1] 0.2
```

As we see, the value of the standard deviations are approximately 0.2

Finally plot the histogram of a large simulations to show that it is close to a normal distribution

```
mns = NULL
for (i in 1 : 1000) mns = c(mns, mean(rexp(40)))
hist(mns, main = 'Histogram for 40 simulations')
```