Statistical Inference Project

Mateus Melo

14/09/2020

Project Overview

This report is divided in two parts. In the first part, we are going to investigate the exponential distribution. Simulating the mean of the distribution by collecting a large sample of averages, we will be able to see how it compares with the Central Limit Theorem. In the second part, we are going to use the toothgrowth data and we will be making assumptions by using confidence intervals.

Simulation Exercise

We start our simulation by collecting a sample of a thousand of exponential averages and storing them in a vector called *exp_avg*. For doing that, we are going to use the *rexp* function, which takes two arguments. The first argument is *n* and it represents the number of random values of the distribution the funciton will return. The second argument is *lambda* which is a rate parameter of the exponential. In this exercise, we will set *lambda* to 0.2 and *n* to 40.

```
exp_avg <- vector("numeric",1000)
for(i in 1:1000){
        exp_avg[i]<-sum(rexp(40,0.2))/40
}
```

The exponential mean and the standard deviation are equal to 1/lambda. Since lambda is 0.2, we will have mean and standard deviation equals to 5. Therefore, the variance will be 25. Let us compare these values to the sample ones.

```
exp_mean<-exp_sd<-5
exp_var<-25
exp_sample_mean<-mean(exp_avg)
exp_sample_sd<-sd(exp_avg)
exp_sample_var<-var(exp_avg)
cat("Difference of means: ",exp_mean-exp_sample_mean)
```

```
## Difference of means:  -0.04285212
```

```
cat("Difference of standard deviations: ",exp_sd-exp_sample_sd)
```

```
## Difference of standard deviations:  4.182038
```

```r
cat("Difference of variances: ",exp_var-exp_sample_var)
```

```
## Difference of variances:  24.33094
```