Statistical Inference

Central Limit Theorem review and some inference about teeth grow

MATIAS, V. A.

05/04/2021

Overview

The first part of the report makes one thousand simulations of the exponential function to analyzes the Central Limit Theorem application under the means distribution. The second part is focused to applies statistic inference techniques bringing a better understanding of teeth growth.

1. Central Limit Theorem

The Central Limit Theorem (CLT) says that the averages of any distribution tend to the normal as the amount of data grows. The first topic examines this fact.

1.1 Simulation

We'll create a thousand exponential distributions and extract the means to build a new distribution that should be next to the normal. The exponential distribution is calculated as:

$f(x) = \lambda e^{-\lambda x}$

Where:

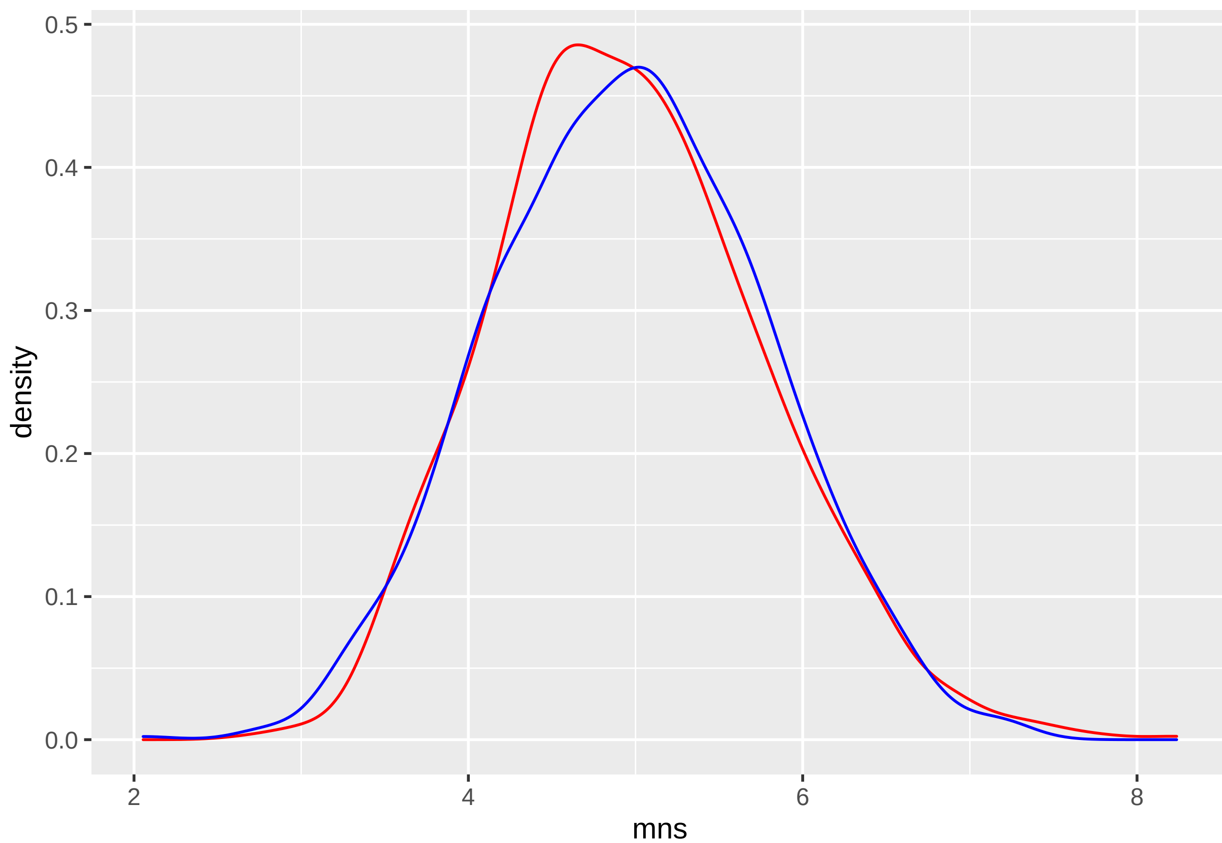
- λ = mean/variance of the data, being that both variance and mean are the same in a Poisson Distribution
- x = The numeric data related to the observations

In this report will be created 1000 of these distributions. Each of them has 40 random values, and lambda equals 0.2, so the formula can be adapted to $f(x) = 0.2e^{-0.2x}$. The code chunk below extracts the means of each of the exponential distributions, store them and plot the means distribution plot.

```
library(ggplot2)

mns <- c()
for (i in seq(1000)) {
  mns <- c(mns, mean(rexp(n = 40, rate = 0.2)))
}

ggplot() +
  geom_density(aes(x = mns), col = "red") +
  geom_density(col = 'blue',aes(rnorm(1000, mean = mean(mns), sd = sd(mns))))
```




```r
library(dplyr)
testt <- t.test(mns, conf.level = 0.95)$conf.int
incon <- mean(mns) + c(1,-1) * qnorm(0.025) * sd(mns)/sqrt(length(mns))

length(mns[mns > testt[1] & mns < testt[2]])
```

```
## [1] 53
```

Sample Mean versus Theoretical Mean

Sample Variance versus Theoretical Variance

Distribution

Teeth

Data

Inference

Conclusions