Simulating the Central Limit Theorem with Exponential Distribution

Bharat Dogra

2025-03-24

Overview

This project investigates the Central Limit Theorem (CLT) using simulations with the exponential distribution. We simulate averages of 40 exponential variables over 1000 trials and compare the resulting sampling distribution with theoretical expectations. We analyze its mean, variance, and shape to confirm convergence to normality. We also conduct basic inferential analysis using the ToothGrowth dataset. ## Simulations We simulate 1000 averages of 40 exponential random variables with a rate parameter `lambda = 0.2`. The theoretical mean and standard deviation of the exponential distribution are both `1 / lambda = 5`.

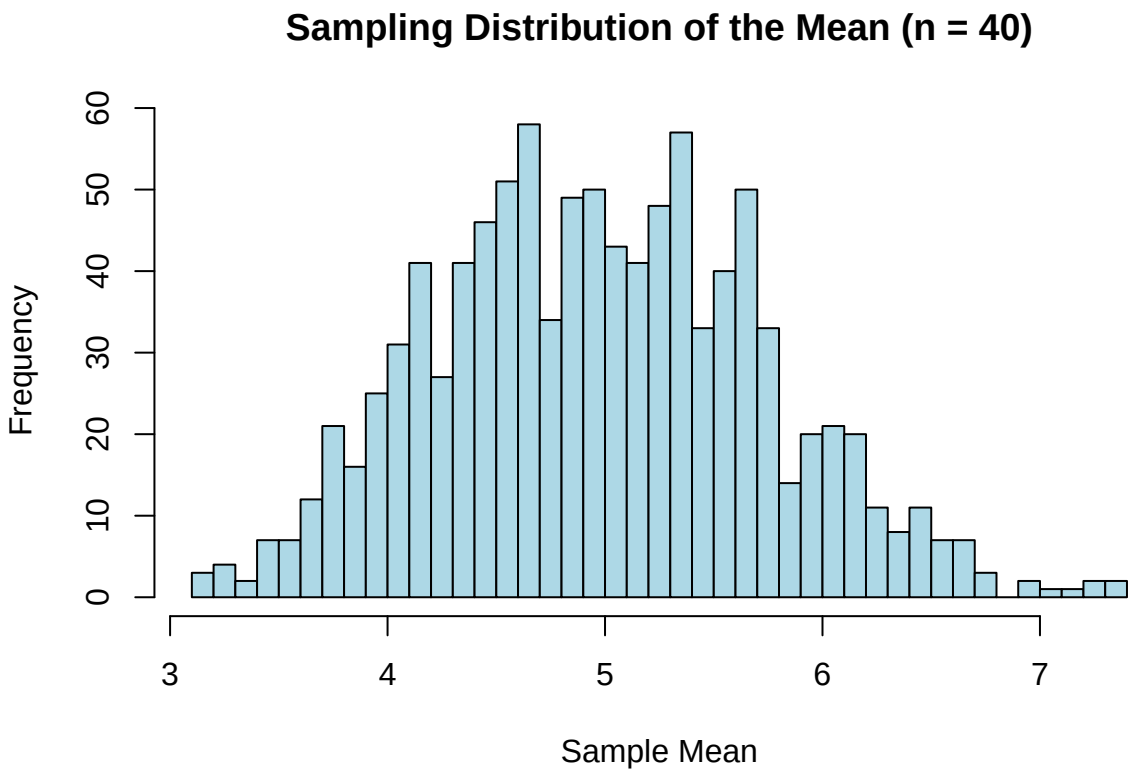
```
set.seed(1234)
lambda <- 0.2
n <- 40
num_sim <- 1000

# Simulate sample means
means <- replicate(num_sim, mean(rexp(n, lambda)))
```

Visualizing the Sampling Distribution

We create a histogram of the sample means and compare the sample mean to the theoretical mean.

```
# Histogram of sample means
hist(means, main = "Sampling Distribution of the Mean (n = 40)",
     xlab = "Sample Mean", col = "lightblue", breaks = 40)
```




```r
# Compare sample mean and theoretical mean
sample_mean <- mean(means)
theoretical_mean <- 1 / lambda
```

Sample Variance vs Theoretical Variance

We now compare the sample variance of the means to the theoretical variance of the sampling distribution of the mean. The theoretical variance is calculated as `(1 / lambda)^2 / n`.

```r
# Calculate sample and theoretical variance
sample_variance <- var(means)
theoretical_variance <- (1 / lambda)^2 / n # Variance of the sample mean

# Output both values
sample_variance
```

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

```r
theoretical_variance
```

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