Statistical Inference Project 1

Robert Tuck

June 20, 2015

Overview

This project is an investigation of the exponential distribution in R and comparison 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. The investigation uses the distribution of averages of 40 exponentials over a thousand simuations.

Simulations

```
number_of_simulations <- 1000
lambda <- 0.2
n <- 40

set.seed(316)

simulations <- matrix(rexp(number_of_simulations * n, rate=lambda), number_of_simulations, n)
simulations_mean <- rowMeans(simulations)
```

Sample Mean versus Theoretical Mean

```
data_mean <- mean(simulations_mean)
data_mean
```

```
## [1] 5.004434
```

```
theoretical_mean <- 1/lambda
theoretical_mean
```

```
## [1] 5
```

Sample Variance versus Theoretical Variance

```
data_variance <- var(simulations_mean)
data_variance
```

```
## [1] 0.6204385
```

```r
theoretical_var <- ((1/lambda)*(1/sqrt(n)))^2
theoretical_var
```

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

Distribution

```r
library(ggplot2)
normal_plot_data <- data.frame(simulations_mean)
g <- ggplot(normal_plot_data, aes(x = simulations_mean))
g <- g + geom_histogram(alpha = .20, colour = "black", aes(y = ..density..))
g <- g + geom_density(color = "red")
```

```r
actual_confidence_interval <- mean(simulations_mean) + c(-1, 1) * qnorm(0.975) * sd(simulations_mean)/s
actual_confidence_interval
```

```
## [1] 4.760334 5.248534
```

```r
theoretical_confidence_interval <- theoretical_mean + c(-1,1) * qnorm(0.975)*sqrt(theoretical_var)/sqrt
theoretical_confidence_interval
```

```
## [1] 4.755005 5.244995
```

```r
qqnorm(simulations_mean)
qqline(simulations_mean)
```