Simulation of Exponential Distribution

Jiachang (Ernest) Xu

6/12/2017

Overview

This project investigates the exponential distribution in R, and compares it with the Central Limit Theorem.

Simulation

This project starts with generating 1000 simulations of 40-size samples of exponential distribution, using the for loop and the rexp() function in R. Each simulation, and its sample mean and its sample standard deviation is stored in respective data frames.

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

data <- NULL
means <- NULL
stdvs <- NULL

for (i in 1:s) {
    set.seed(i)
    data <- c(data, rexp(n, lambda))
    means <- c(means, mean(rexp(n, lambda)))
    stdvs <- c(stdvs, sd(rexp(n, lambda)))
}
```

Sample Mean versus Theoretical Mean

Using the simulation data to compare sample mean with theoretical mean

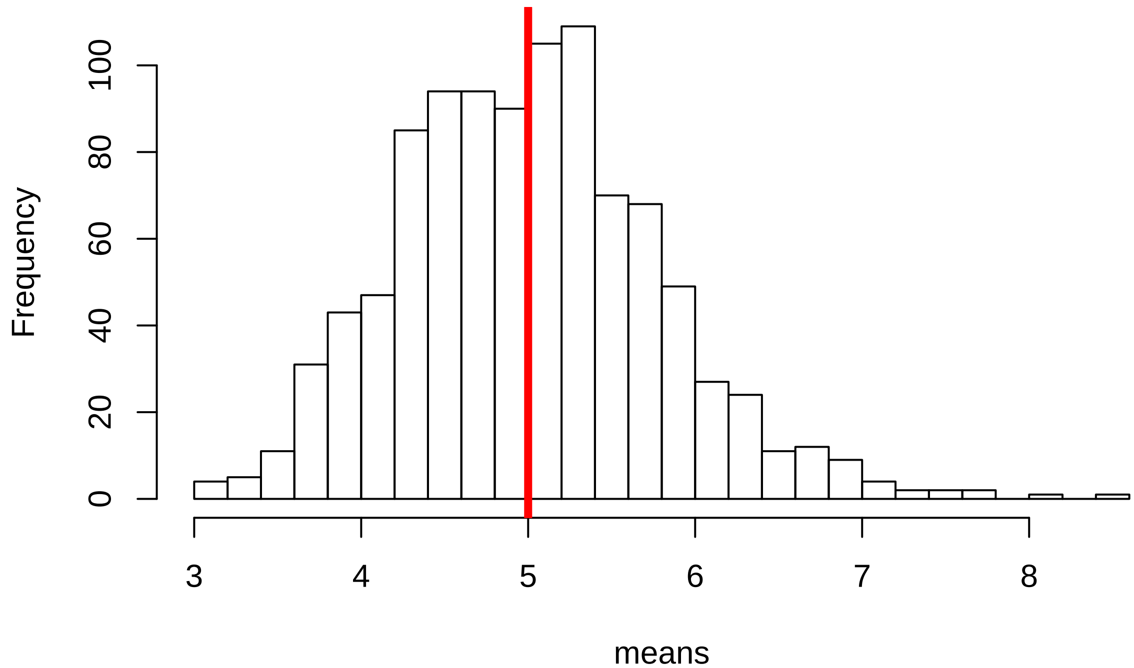
```
sample_mean <- mean(means)
sample_mean
```

```
## [1] 5.011243
```

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

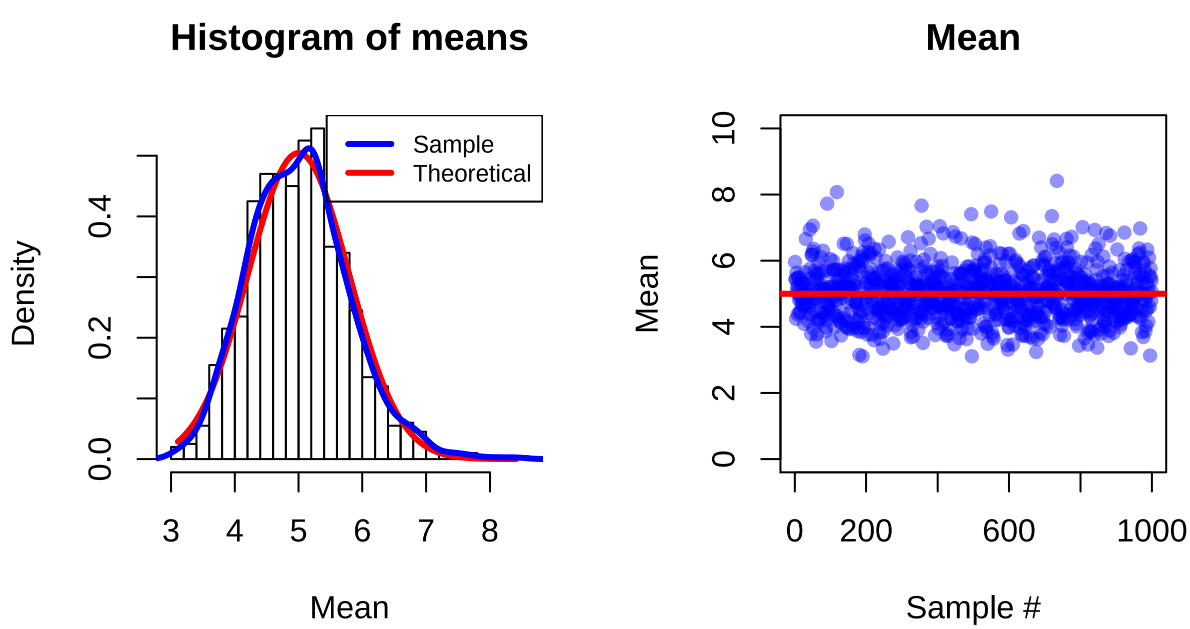
```
## [1] 5
```

Figure 1: Sample Mean vs. Theoretical Mean



The sample mean deviates from the theoretical mean by only 0.011. Figure 1 shows the sampling distribution is approximately normal, and centered very close to the theoretical mean (red line).

Figure 2: Distribution of Sample Means



Sample Variance versus Theoretical Variance

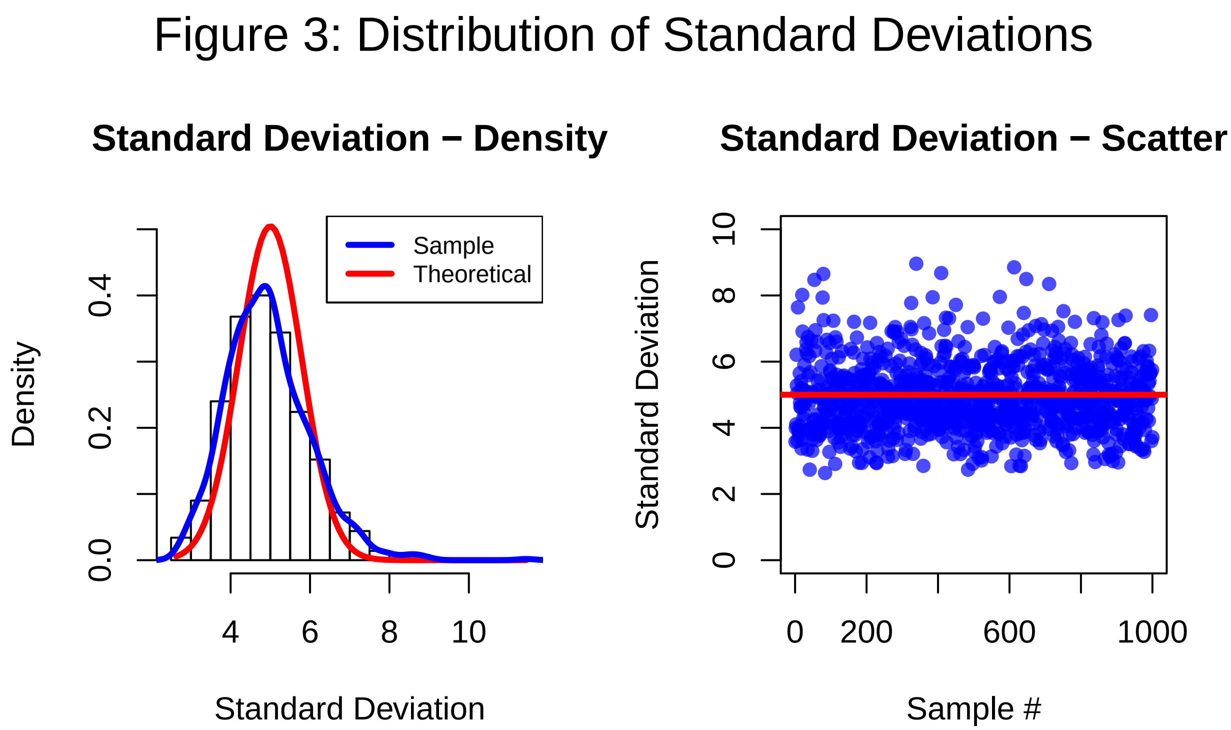
```
sample_var <- var(means)
sample_var
```

```
## [1] 0.60981
```

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

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

The sample and theoretical variances are very close



Appendix I: Figure 1 Plotting Code

```
hist(means, 22, main = "Figure 1: Sample Mean vs. Theoretical Mean")
abline(v = theoretical_mean, lw = 4, col = "red")
```

Appendix II: Figure 2 Plotting Code

```
par(mfrow=c(1,2), oma=c(2,0,2,0))

# histogram as density function
hist(means, 22, freq = FALSE, xlab = "Mean", ylab = "Density")

# the theoretical line
xfit <- seq(min(means), max(means), length = 100)
yfit <- dnorm(xfit, mean = 1/lambda, sd = 1/lambda/sqrt(n))
```

```
lines(xfit, yfit, pch=22, lty = 1, lw = 3, col = "red")

# the sample line
lines(density(means), lw = 3, col = "blue")

# the title and legend
mtext("Figure 2: Distribution of Sample Means", outer = TRUE, cex = 1.5)
legend('topright', c("Sample", "Theoretical"), col=c("blue", "red"), lw=c(3,3), cex = .75)

#plotting the data on the mean for the 1000 simulations
plot(means, main = "Mean", ylab = "Mean", xlim = c(0,s), ylim = c(0,10), xlab = "Sample #", pch = 16, c
abline(h=5, col = "red", lw =3)
```

Appendix III: Figure 3 Plotting Code

```
par(mfrow=c(1,2), oma=c(2,0,2,0))
hist(stdvs, 22, freq = FALSE, ylim = c(0,.5), main = "Standard Deviation - Density", xlab = "Standard De

# the theoretical line
xfit <- seq(min(stdvs), max(stdvs), length = 100)
yfit <- dnorm(xfit, mean = 1/lambda, sd = 1/lambda/sqrt(n))
lines(xfit, yfit, pch=22, lty = 1, lw = 3, col = "red")

# the sample line
lines(density(stdvs), lw = 3, col = "blue")

# the title and legend
mtext("Figure 3: Distribution of Standard Deviations", outer = TRUE, cex = 1.5)
legend('topright', c("Sample", "Theoretical"), col=c("blue", "red"), lw=c(3,3), cex = .75)

plot(stdvs, main = "Standard Deviation - Scatter", ylab = "Standard Deviation", xlim = c(0,s), ylim = c
abline(h=5, col = "red", lw =3)
```