Inference Project4: Simulation Exercise

Shuai Wang

Dec, 11, 2016

Part 1: Simulation Exercise

Title:

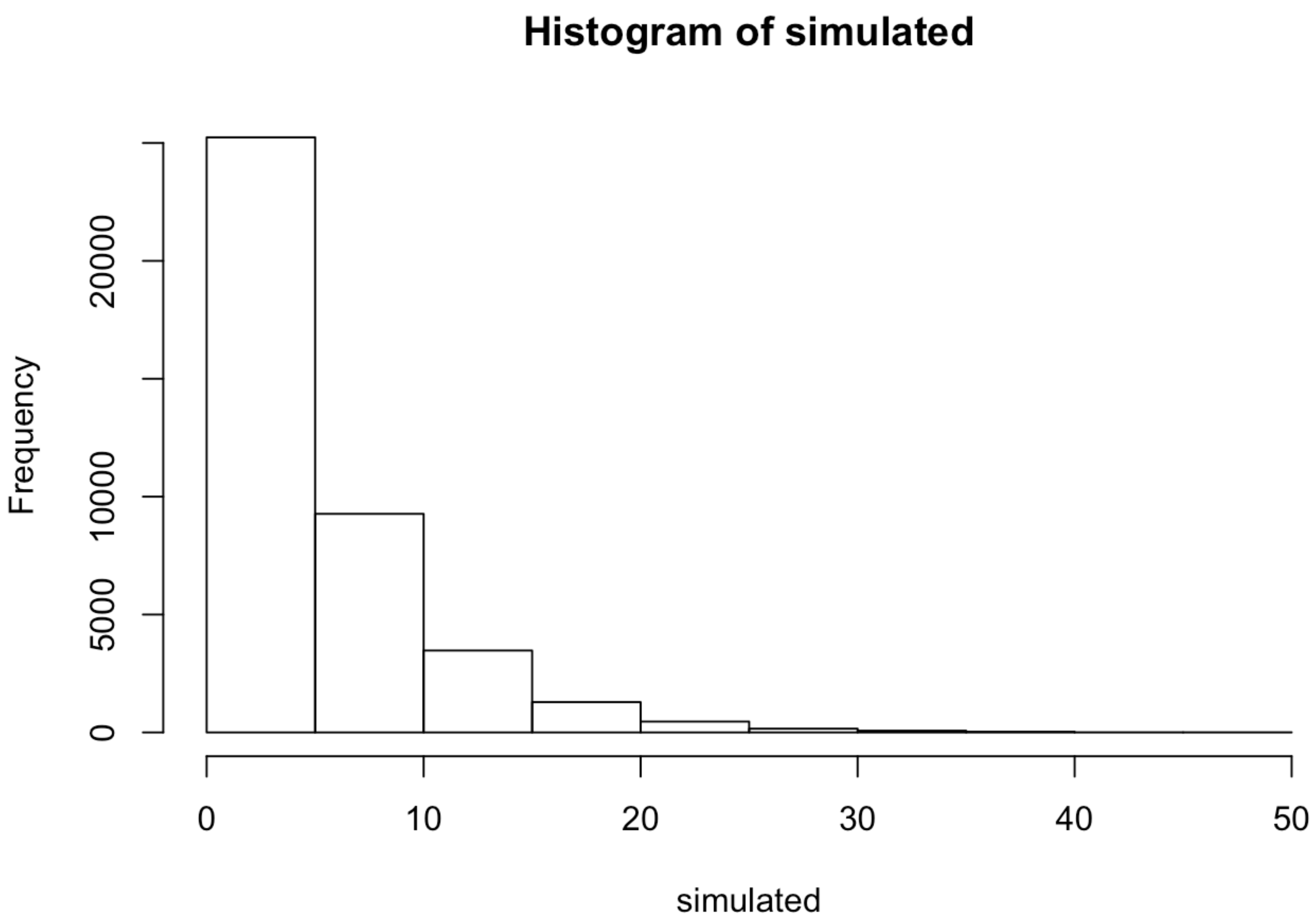
Comparing Simulation Results of exponential distribution with Central Limit Theorem

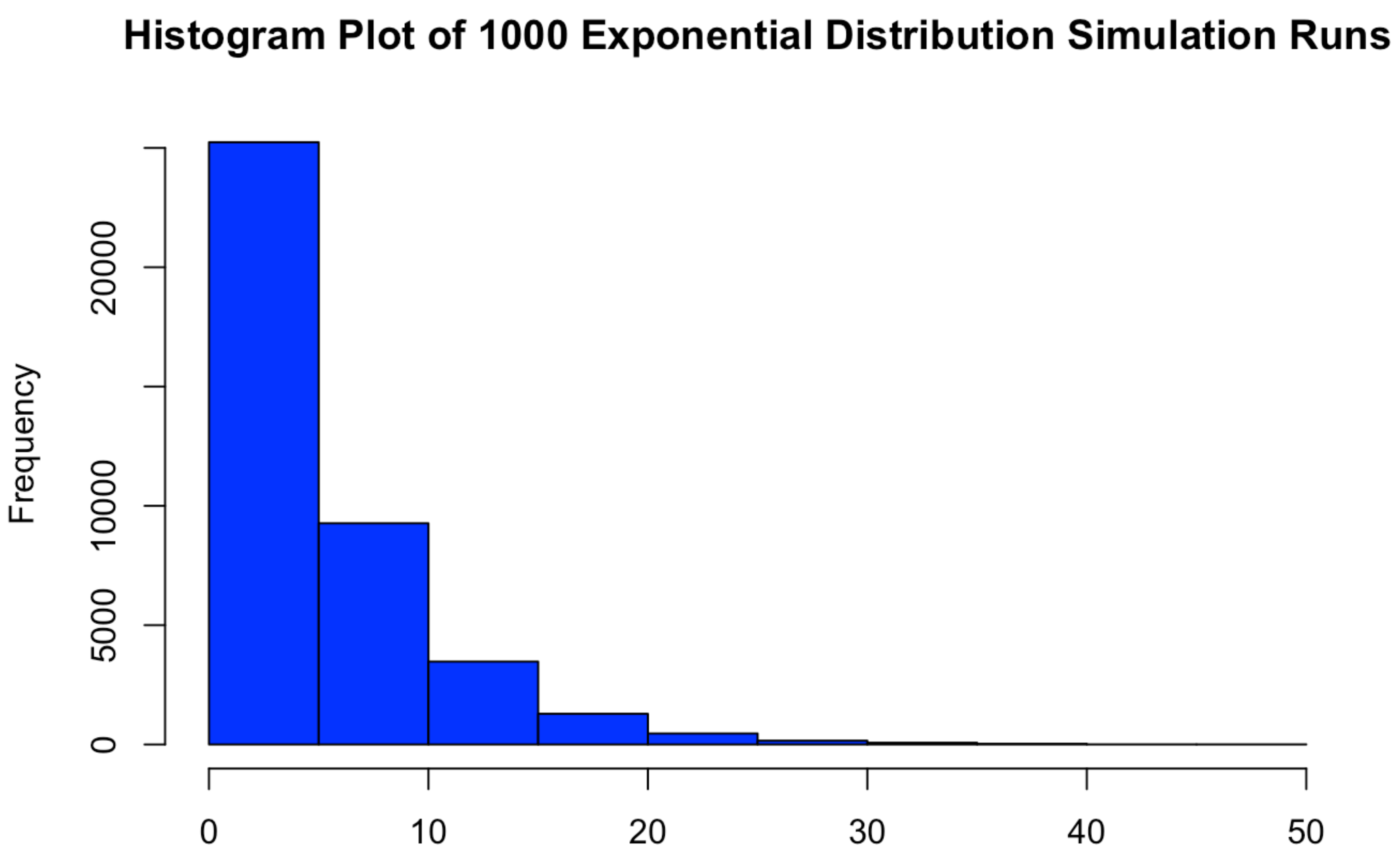
Overview:

In the first part of the project, I will run simulation results of exponential distribution. The exponential distribution number generator is based on rexp() function.

Simulations:

```r
# setting up the experiment parameters
nosim <- 1000
n <- 40
lambda <- 0.2
simulated<-matrix(rexp(nosim * n, lambda), nosim)
plot(hist(simulated),xlab="Value",ylab="Frequency",col="blue",main="Histogram Plot of
1000 Exponential Distribution Simulation Runs")
```


Histogram of simulated
Frequency
0
5000
10000
20000
0
10
20
30
40
50
simulated


Histogram Plot of 1000 Exponential Distribution Simulation Runs
Frequency
0
5000
10000
20000
0
10
20
30
40
50

Value

From the instruction, we know that the the mean of exponential distribution is 1/lambda and the standard deviation is also 1/lambda.

Sample Mean versus Theoretical Mean:

Include figures with titles. In the figures, highlight the means you are comparing. Include text that explains the figures and what is shown on them, and provides appropriate numbers.

Here in the figure below. I present the distribution of the normalized means of 1000 simulations in blue histogram plot. I also plot standard normal distribution in the black curve.

```
#plot(hist_plot,xlab="Value",ylab="Frequency",col="blue",main="Histogram Plot of One Exponential Distribution Simulation Run")
observed.mean<-apply(simulated, 1, mean)
normalizedMean<-observed.mean-1/lambda
normalizedData<-normalizedMean*sqrt(40)/(1/(lambda))
hist(normalizedData,xlab="",ylab="",yaxt='n',xaxt='n',main="",col="blue")
par(new=TRUE)
x<-rnorm(1000, mean = 0, sd = 1)
p<-density(x,main="",xlab="",ylab="")
plot(p,main="",xlab="",ylab="")
title(main="Distribution of Means of Exponential with Normal Distribution Overlap",xlab="value",ylab="Frequency",sub="Figure 1. Simulated Exercise")
```

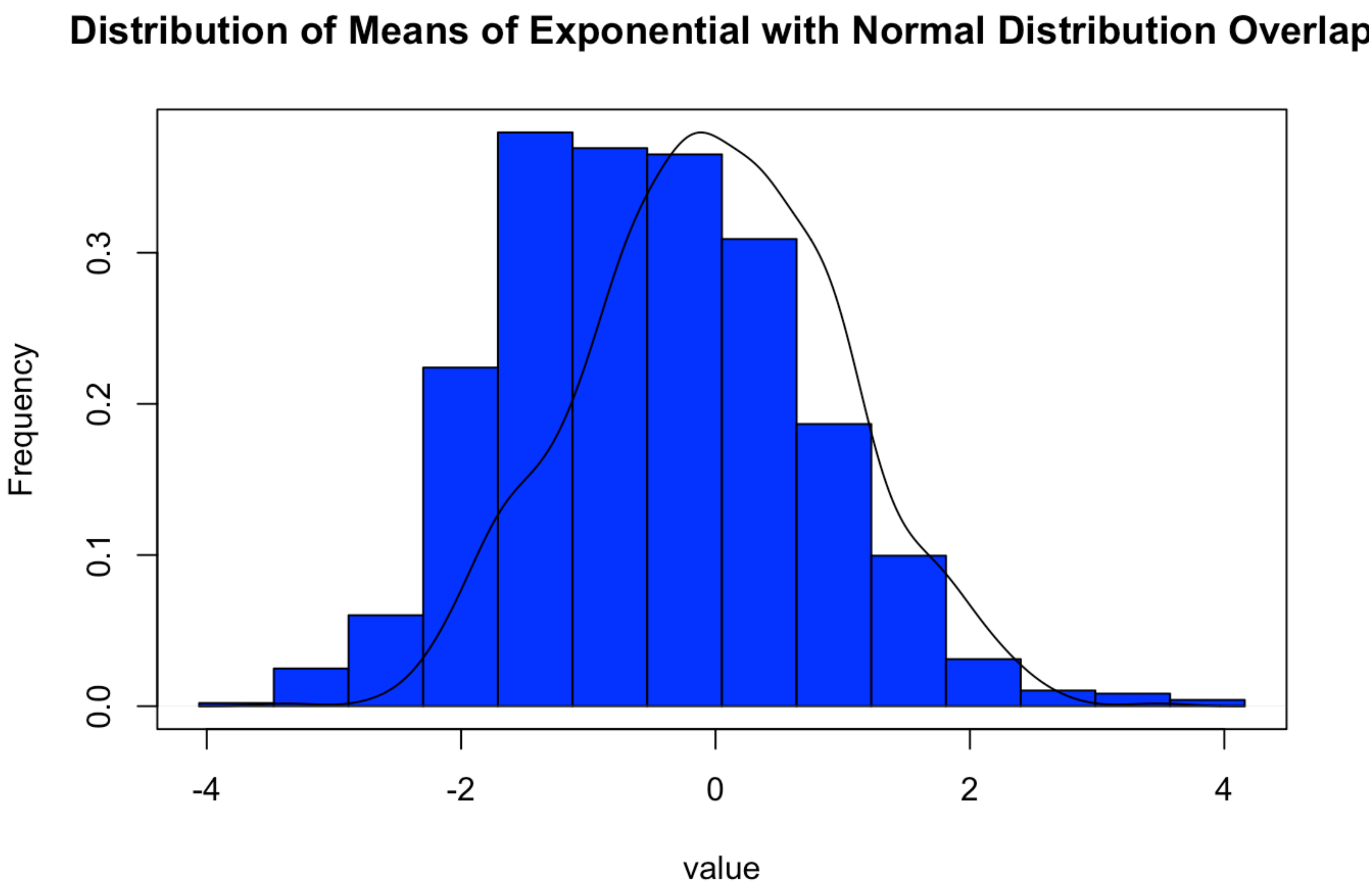


Figure 1. Simulated Exercise

```
#observed.sd
observed.var<-var(apply(simulated,1,mean))
#observed.var
#sqrt(observed.var)
#abline( v = observed.mean, col = "black")
#text(1,0, "Obsevered Mean", col = "white", adj = c(0, -.1))
#abline( v = 1/lambda, col = "red")
#text(3,0, "Theoretical Mean", col = "white", adj = c(0, -.1))
```

Sample Variance versus Theoretical Variance:

Include figures (output from R) with titles. Highlight the variances you are comparing. Include text that explains your understanding of the differences of the variances.

```
observed.var<-var(observed.mean)
sprintf("The observed variance of mean should be:%.2f",observed.var)
```

```
## [1] "The observed variance of mean should be:0.61"
```

```
expected.var<-(1/lambda^2)/n
sprintf("The expected variance of mean should be:%.2f",expected.var)
```

```
## [1] "The expected variance of mean should be:0.62"
```

Distribution:

Via figures and text, explain how one can tell the distribution is approximately normal.

Based on the Central Limit Theorem, we learn that the mean of random distributions follow standard normal distribution when the repetition of simiulations gets larger. From figrue 1, we can see that the distribution of 1000 means of 40 expeontials follow nicely with the standard normal distribution. The observed variance of the mean is also about the same as the expected variance of the mean. From these results, we can tell the distribution is approximately normal.