Statistic Inference Project Assignment Part1

Pete Schnettler

January 13, 2016

Overview

Comparing the simulated mean and variance with the theoretical values

We will run 1000 rounds of simulation of 40 exponentials with $\lambda = 0.2$, using a fixed seed, and comparing the distribution of the simulated mean and variance with the theoretical value of $1/\lambda$:

```
nSim <- 1000
nVals <- 40
lambda <- 0.2
set.seed(456)

#Create exponential distribution and transpose into a matrix
simMatrix <- t(replicate(nSim, rexp(nVals, lambda)))

#Create a datafrom consisting of the Mean and variance
df <- data.frame(
        Mean=c(mean(rowMeans(simMatrix)), 1/lambda),
        Variance=c(mean(apply(simMatrix, 1, var)), 1/lambda^2))
rownames(df) <- c("Sample", "Theoretical")
library(knitr)
```

```
## Warning: package 'knitr' was built under R version 3.2.3
```

```
kable(df)
```

	Mean	Variance
Sample	4.984081	24.70639
Theoretical	5.000000	25.00000

1.Show the sample mean and compare it to the theoretical mean of the distribution

The Sample mean is 4.984081 while the Theoretical mean is 5.0 while are within a couple of decimal points, as expected by the Central Limit Theorem.

2.Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution

The sample variance 24.7 and theoretical variance 25 are also close

# 3.Show that the distribution is approximately normal

Also, according to the CLT, the distribution of the simulated means should be approximately normal. .

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.2.2
```

```
#Calculate the row means
meanVals <- rowMeans(simMatrix)


zMean <- (meanVals - mean(meanVals)) / sd(meanVals)
qplot(zMean, geom = "blank") +
  stat_function(fun = dnorm, aes(colour = 'Normal')) +
  geom_histogram(aes(y = ..density..), alpha = 0.4, binwidth=.35) +
  geom_vline(xintercept=0, colour="red", linetype="longdash") +
  scale_colour_manual(name = 'Density', values = c('red', 'blue')) +
  ylab("Density") + xlab("z") +
  ggtitle("Distribution of means of 40 Exponential Distribution Lambda of 0.
2") +
  theme_bw() +
  theme(legend.position = c(0.85, 0.85))
```

Distribution of means of 40 Exponential Distribution Lambda of 0.2
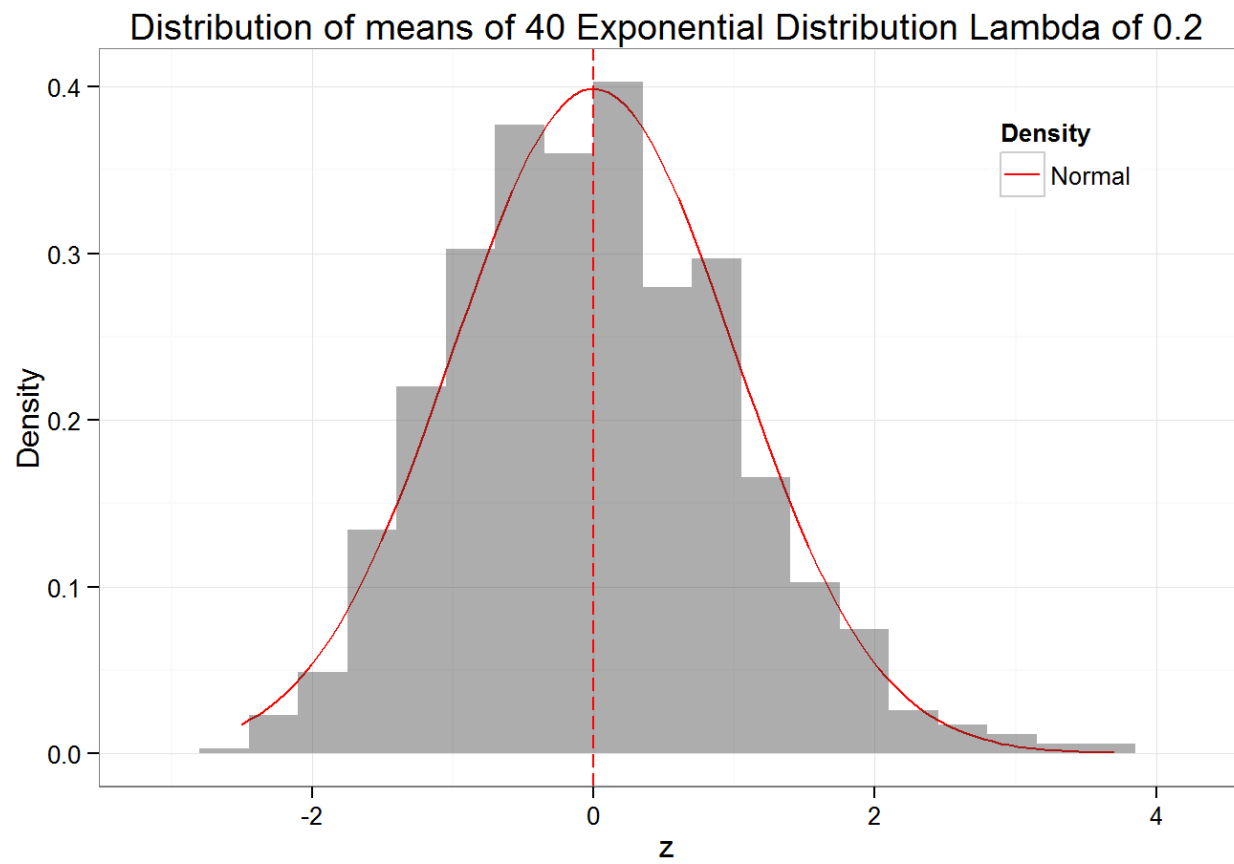
The plot of the means above appears approximately normal