Exponential sample distribution comparison

Cliff Lau

Overview

We will investigate the exponential distribution and compare it with the Central Limit Theorem. We will investigate the distribution of averages of 40 exponentials, each with 1000 simulation.

More specifically, we will be comparing the following:

1. Sample mean versus theoretical mean
2. Sample variance versus theoretical variance
3. Distribution of the sample mean

Simulation

We be using the following simulate for our analysis:

```
library(ggplot2)
library(dplyr)

n <- 40
nosim <- 1000
lambda <- 0.2
set.seed(4321)

sim <- function(nosim, n, lambda) {
  mns <- NULL
  for(i in 1:nosim) {
    mns <- c(mns, mean(rexp(n, lambda)))
  }
  tbl_df(data.frame(average = mns))
}

sim1 <- sim(nosim, n, lambda)
```

Sample Mean versus Theoretical Mean

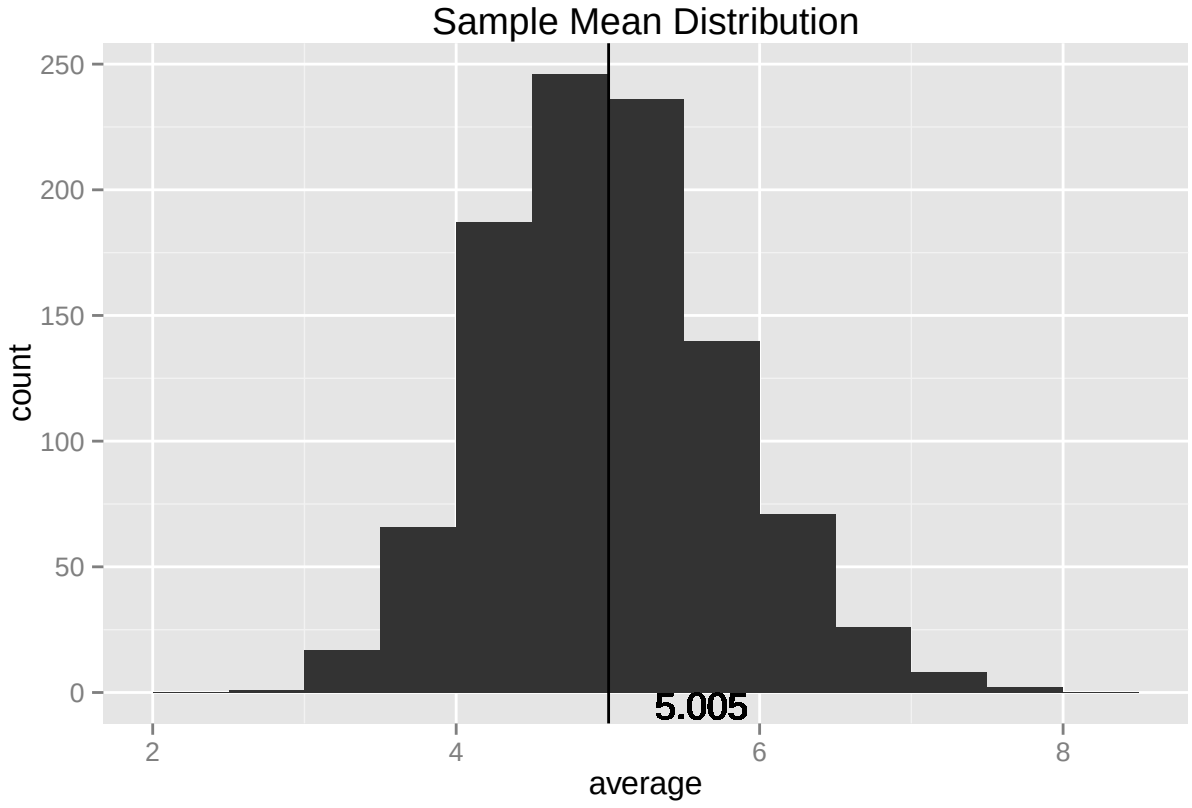
To get the sample and theoretical mean:

```
# Simulation mean
simmean <- mean(sim1[[1]])
# Theoratical mean
theomean <- 1/0.2
```

Sample Mean	Theoretical Mean
5.0045739	5

As you can see from the table above, the sample mean is very close to the theoretical mean. Furthermore is a histogram of the sample mean distribution:

```
ggplot(sim1, aes(average)) +
  geom_histogram(binwidth = 0.5) +
  geom_vline(x = simmean) +
  geom_text(x = simmean, y = 0, hjust = -0.5, vjust = 1,
            label = round(simmean,3)) +
  ggtitle("Sample Mean Distribution")
```



Sample Variance versus Theoretical Variance

To get the sample and theoretical mean:

```
# Simulation variance
simvar <- var(sim1[[1]])
# Theoratical variance
theovar <- 1/0.2
```

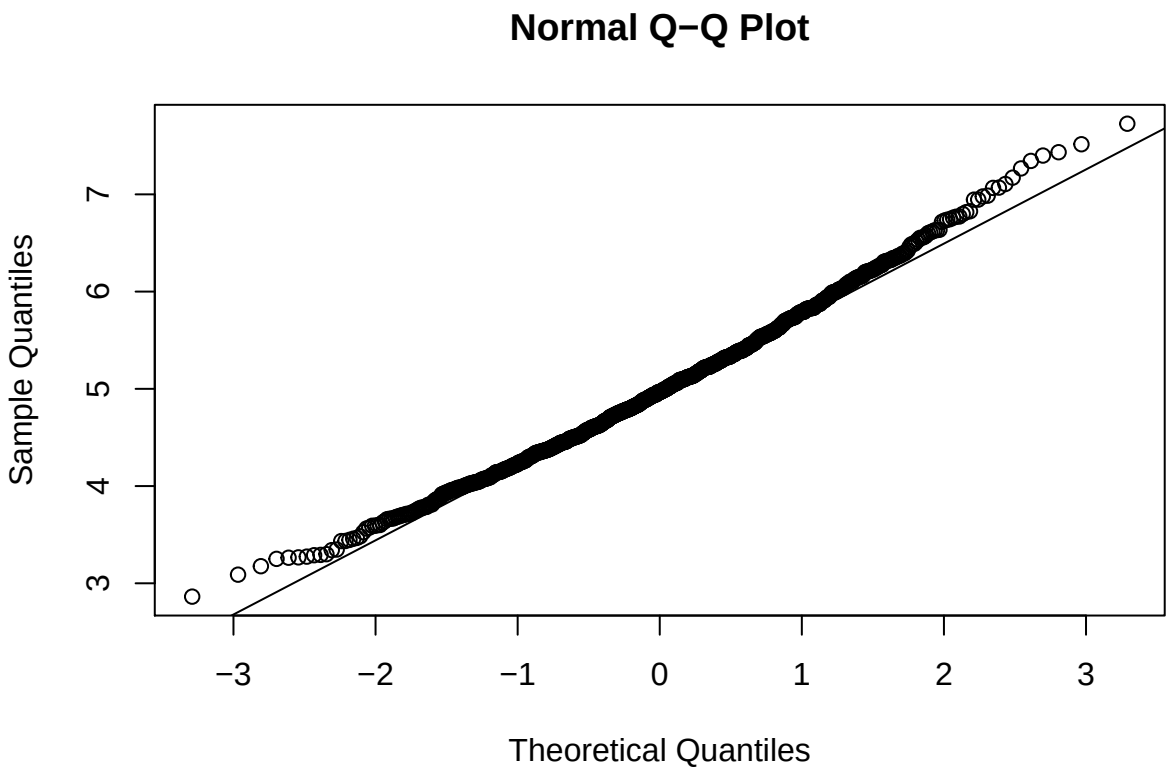
Sample Variance	Theoretical Variance
0.6021698	5

As you can see from the table above, the sample variance is **not** the same as the theoretical variance of the exponential distribution.

Distribution

As we shown above, the distribution of the sample mean has a different variance than the underlying variance of the exponential distribution. Furthermore, we can show that the sample mean distribution is normal by looking at the normal QQ plot:

```
qqnorm(sim1[[1]])
qqline(sim1[[1]])
```



As you can see from the plot above, the sample distribution has similar quantile as normal distribution.