

Exponential Distribution Simulation

mickfenneck

Introduction

The project is meant to answer to this particular question:

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. In this simulation, you will investigate the distribution of averages of 40 exponential(0.2)s. Note that you will need to do a thousand or so simulated averages of 40 exponentials.

Illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponential(0.2)s. You should:

1. Show where the distribution is centered at and compare it to the theoretical center of the distribution.
2. Show how variable it is and compare it to the theoretical variance of the distribution.
3. Show that the distribution is approximately normal.
4. Evaluate the coverage of the confidence interval for $1/\lambda$:: $\bar{X} \pm 1.96 \frac{S}{\sqrt{n}}$.

The Solution

Generating the Data

In order to start simulating the distribution I set the seed so everybody should reproduce in the same way this experiment. The text says to set the number of tests equal to 1000 and λ equal to 0.2

```
set.seed(1994)
lambda <- 0.2
numTests <- 1000
testCount <- 40
data <- matrix(rexp(numTests * testCount, rate=lambda), nrow = numTests)
means <- rowMeans(data)

# Compute the mean of each row of the generated data.
meanDist <- apply(data, 1, mean)
```

We can now concentrate on some particular values:

```
centre <- round(mean(meanDist), 3)
theoreticalmean <- round(1/lambda, 3)

SD <- round(sd(meanDist), 3)
theoreticalSD <- round(1/(lambda * sqrt(testCount)), 3)

var <- round(var(meanDist), 3)
theoreticalvar <- round((1/(lambda * sqrt(testCount)))^2, 3)
```

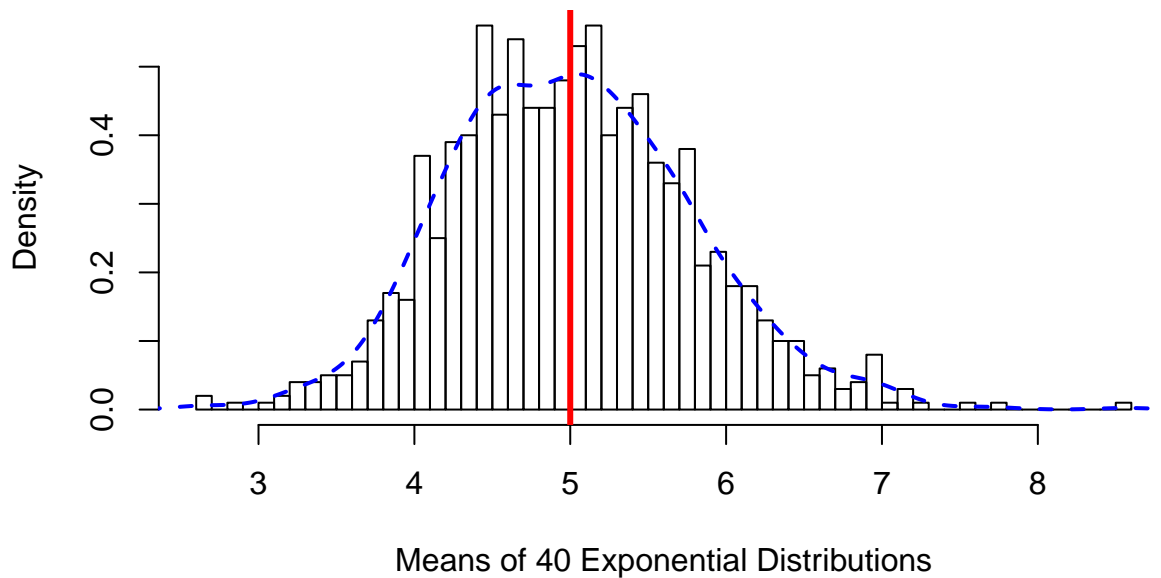
After the computation of the values we can quickly compare them:

- mean: $|\text{centre-theoreticalmean}| = 0.012$
- standar deviation: $|\text{SD-theoreticalSD}| = 0.011$
- variance: $|\text{var-theoreticalvar}| = 0.017$

As we can see, the values are not only comparable but almost the same. We can quickly provide a graphical rappresentation by plotting the data.

```
# plot the histogram of the means
hist(means, breaks=50, prob=TRUE,
     main="Distribution of means of 1000 samples having a lambda of 0.2",
     xlab="Means of 40 Exponential Distributions",
     ylab="Density")
lines(density(means), col="blue", lwd=2, lty=2)
# show the mean of distribution
abline(v=1/lambda, col="red", lwd=3)
```

Distribution of means of 1000 samples having a lambda of 0.2

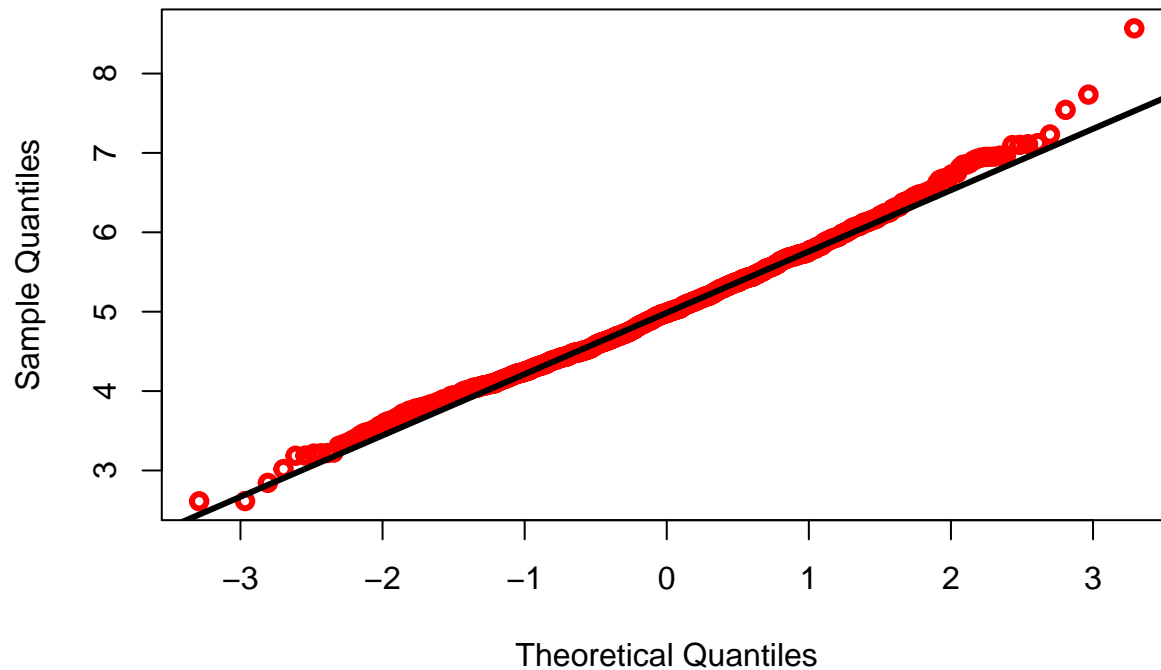


The plot shows that the distribution get close to the normal distribution.

Using a QQplot we can show that the distribution is very similar to the normal, apart from the tails that tends to diverge from the line.

```
qqnorm(means, col="red", lwd=3)
qqline(means, col="black", lwd=3)
```

Normal Q-Q Plot



Final consideration

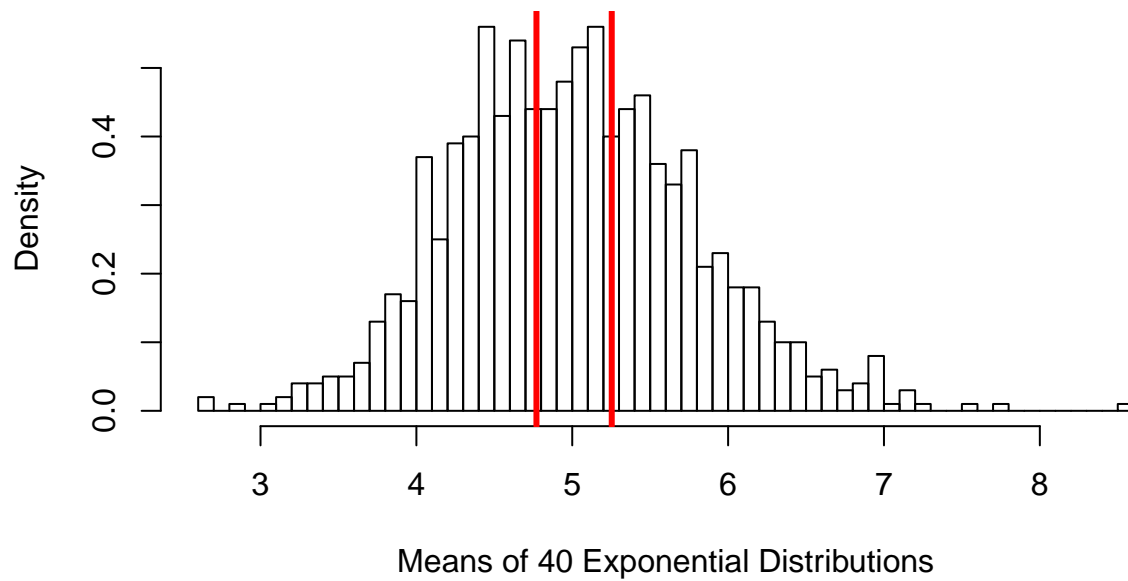
We can finally compute the 95% confidence intervals:

```
CI <- 1.96
ll <- mean(means) - CI * (sd(means)/sqrt(testCount))
ul <- mean(means) + CI * (sd(means)/sqrt(testCount))
```

Lower interval 4.77 and upper interval 5.253.

```
# plot the histogram of the means
hist(means, breaks=50, prob=TRUE,
      main="Distribution of means of 1000 samples having a lambda of 0.2",
      xlab="Means of 40 Exponential Distributions",
      ylab="Density")
# show the two CIs
abline(v=c(ll, ul), col="red", lwd=3)
```

Distribution of means of 1000 samples having a lambda of 0.2



Code Reproducibility

You can reproduce the code how many times you want. The R source is available on github.com at <https://github.com/mickfenneck/statinference>.