Initial Statistical Inference on 40 Exponents

Danish Tamboli

8/5/2020

Initial Statistical Inference on 40 Exponents, and comparison with CLT Distribution.

Overview:

- Comparison between Exponential Distribution and Central Limit Theorem Distribution (Normal Distribution) of 40.

What does Central Limit Theorem state?

- The central limit theorem states that if you have a population with :
  - mean mu.
  - standard deviation sigma.
  - Take sufficiently large random samples from the population with replacement.
  - The distribution of the sample means will be approximately normally distributed.
  - Here we are assuming lambda = 0.2, mu = 1/lambda and sigma = 1/lambda.
  - Here we will be taking 1000 simulations or sample sets from the 40 exponents.

Simulation:

```
set.seed(103101)
init_exp <- rexp(40,0.2)
exp_mean <- mean(init_exp)
print(paste("Observed Exponent Mean:", exp_mean))
```

```
## [1] "Observed Exponent Mean: 5.17300165083614"
```

```
clt <- sapply(1:1000, function(i){mean(sample(init_exp,replace = TRUE))})
clt_mean <- mean(clt)
print(paste("Observed Cental Limit Theorem Mean:", clt_mean))
```

```
## [1] "Observed Cental Limit Theorem Mean: 5.18666396556655"
```

Sample Mean versus Theoretical Mean:

- From the above we can discuss the Theoretical and Observed means:
  - We see that as lambda is 0.2, Hence the Theoretical mean = 1/lambda should be 5
  - The Observed Exponent Mean and Observed Central Limit Theorem Mean is ~ the Theoretical mean.

```
variance_exp <- var(init_exp)
print(paste("Observed Exponent Variance:",variance_exp))
```

```
## [1] "Observed Exponent Variance: 20.4974576001592"
```

```
variance_clt <- var(clt)
print(paste("Observed Cental Limit Theorem Variance:",variance_clt))
```

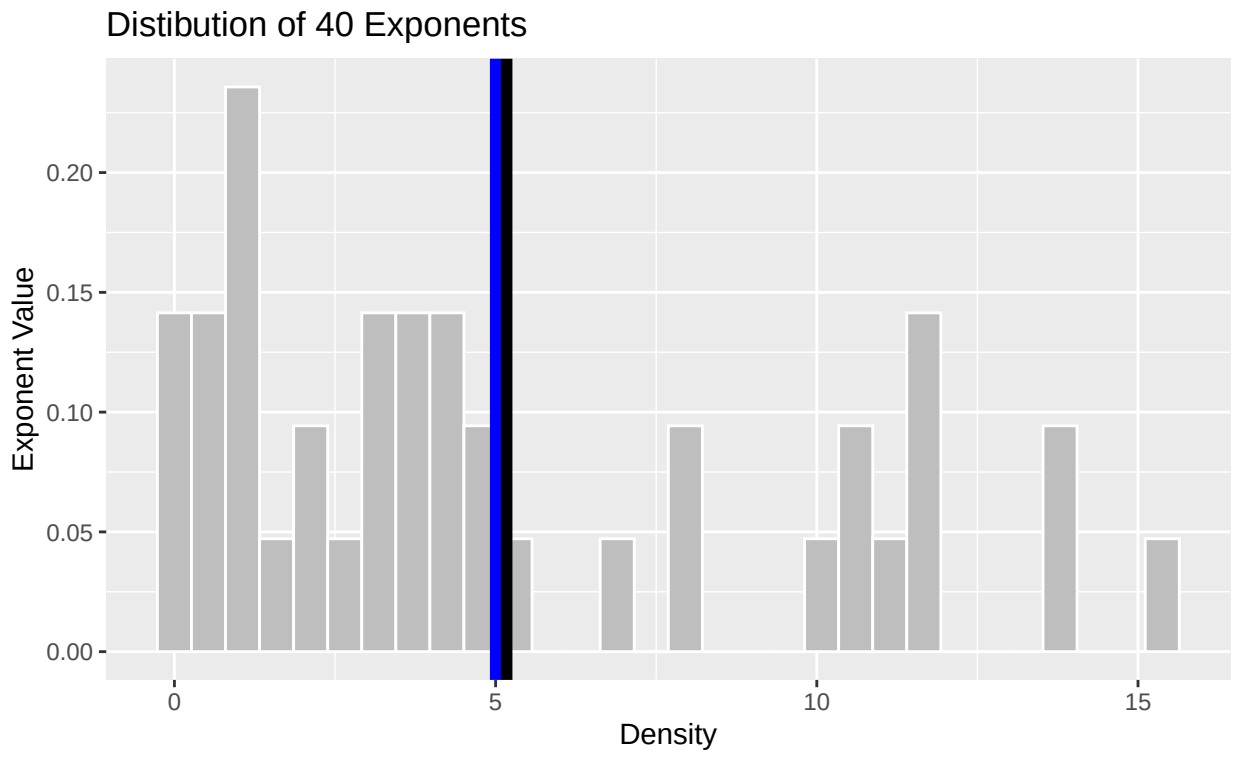
```
## [1] "Observed Cental Limit Theorem Variance: 0.511316406910704"
```

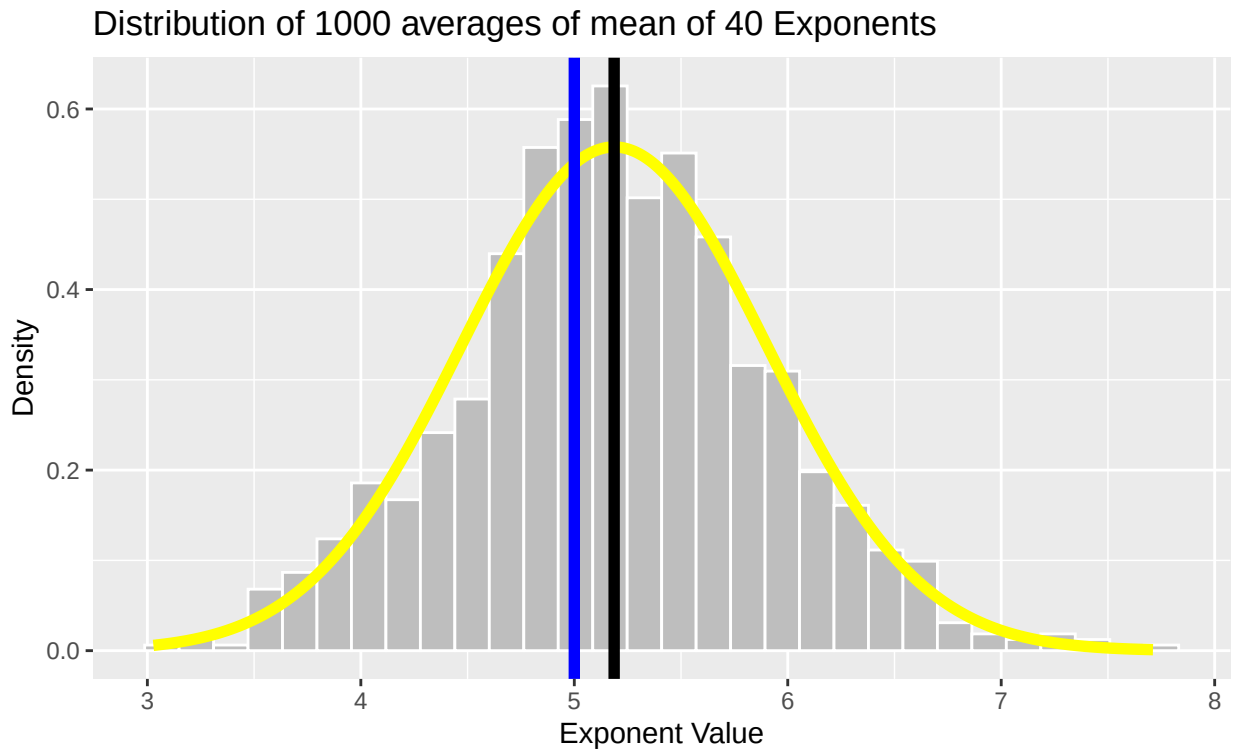
Sample Variance versus Theoretical Variance:

- From the above we can discuss the Theoretical and Observed variances:
  - We see that as lambda is 0.2, Hence the Theoretical variance = sdˆ2, sd = (1/lambda) = 5 should be 25
  - The Observed Exponent Variance is ~ the Theoretical mean.

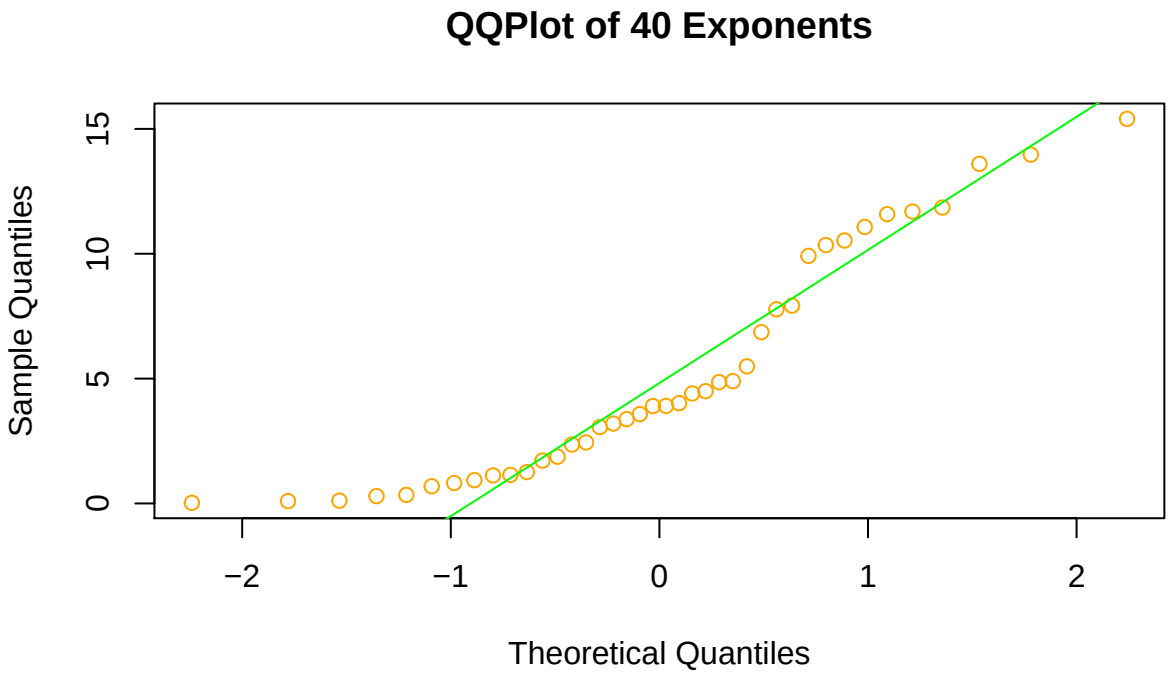
Distribution:

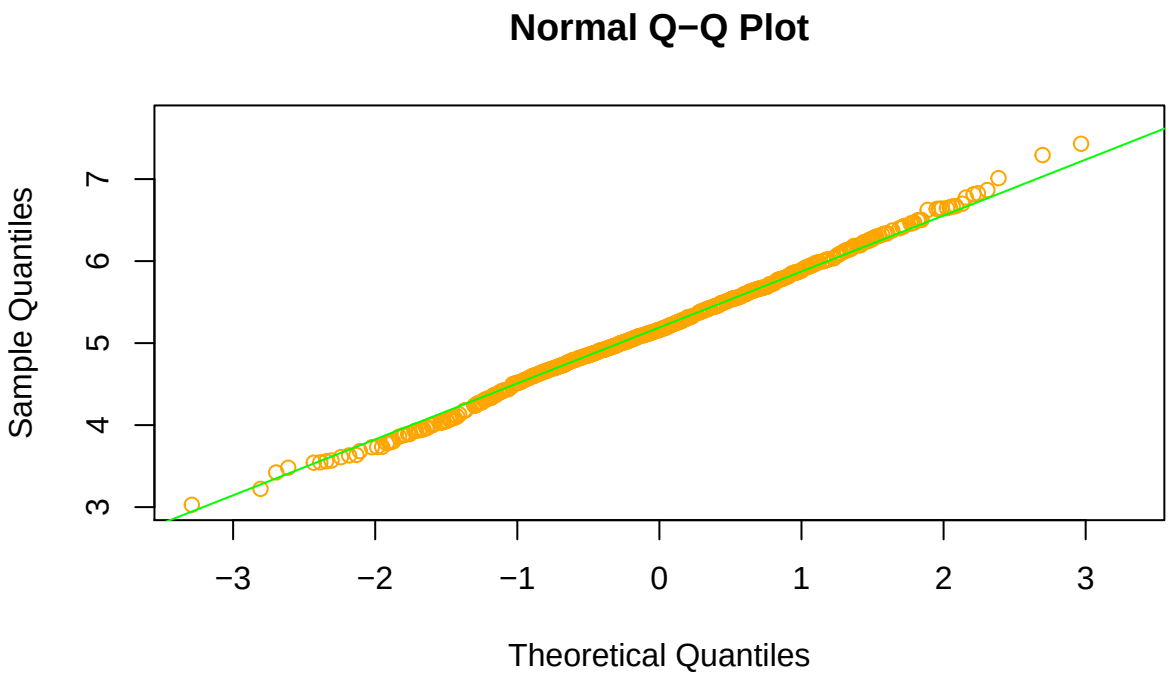
```
## Warning: package 'ggplot2' was built under R version 4.0.2
```





- From the above 2 figures we can see how the Central Limit Distribution (grey bars) is approximately normal (yellow line), as how it fits under the curve of the normal line (yellow).





- From the above 2 figures we can clearly see:
  - How the 40 exponents is not normal and hence the points are not on the straight line.
  - How the CLT Distribution is of Normal Distribution Type as majority of the points are on the straight line.