Exponential Distribution Investigation

Danny Fraser

October 2015

Overview

This paper is an investigation of the exponential distribution, and a comparison of that distribution with the Central Limit Theorem.

The exponential distribution is simulated in R using the `rexp()` function, which takes two parameters: `n` for the number of observations, and `lambda`(λ) for the exponent rate. We will be using $\lambda = 0.2$.

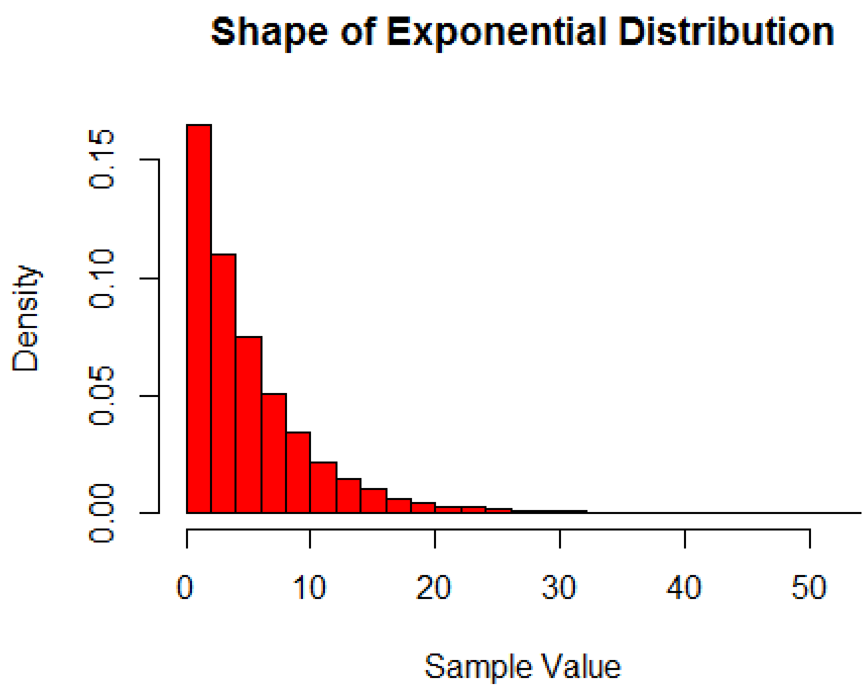
Simulations

Our simulations will be of 1000 averages of 40 random samples from the exponential distribution.

```
sims = 1000
n = 40
lambda = 0.2

samples <- matrix(rexp(n * sims, lambda), sims, n)
sampleMeans <- rowMeans(samples)
```

Properties of the Exponential Distribution



Sample Mean vs Theoretical

One of the characteristics of the exponential distribution is that the mean can be given by $1/\lambda$, so in this case we would expect a theoretical population mean μ to be 5.

We can test this by looking at the distribution of the sample mean, $\overline{X}$, for our 1000 samples of 40 random draws from the distribution:

```
mean(sampleMeans)
```

```
## [1] 4.98487
```

Sample Standard Deviation & Variance vs Theoretical

We also know that the standard deviation σ of the distribution is given by $1/\lambda$, which is again 5 for our parameters.

The Central Limit Theorem tells us that the sample mean standard deviation ('standard error of the mean') is equal to the population standard deviation divided by the square root of the sample size:

$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$ i.e. $\sigma_{\overline{x}} = \frac{1/\lambda}{\sqrt{n}}$

Knowing that the variance is the square of the standard deviation, we can now calculate the theoretical variance σ^2:

```r
theoreticalVar = ((1 / lambda) / sqrt(n)) ^ 2
theoreticalVar
```

```
## [1] 0.625
```

We can calculate the actual variance using R's `var()` function:

```r
var(sampleMeans)
```

```
## [1] 0.6124582
```

We can see that this is close but not exact - a greater number of simulations would be necessary to reduce the difference between these values.

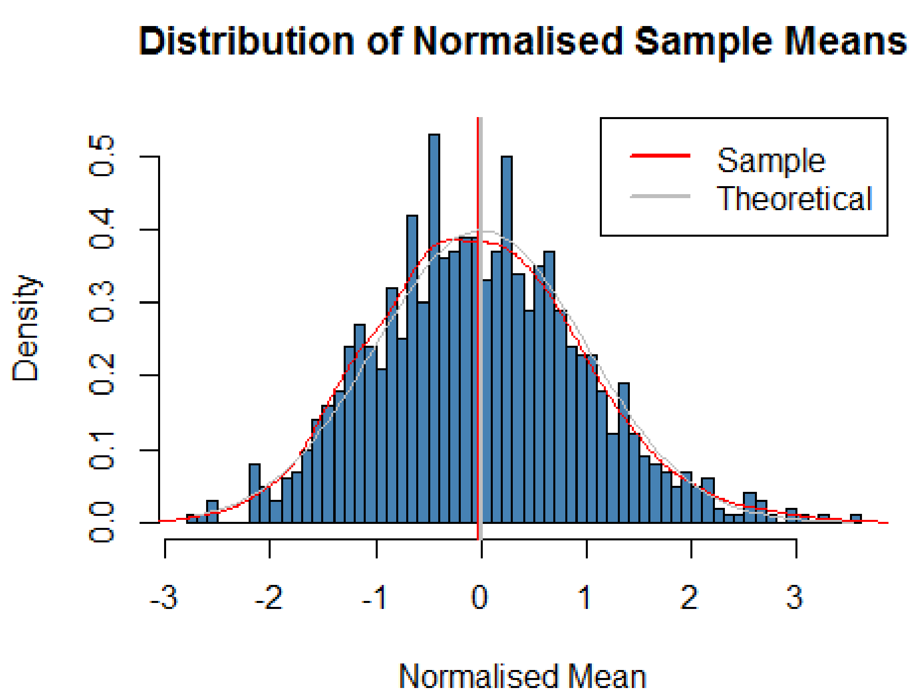
Distribution of Sample Means vs Theoretical

Now we can plot a histogram of normalised sample means to see how they are distributed.

The normalised mean is defined as the difference between the sample and theoretical means divided by the standard error of the sample mean:

$$\frac{\mu_{\bar{x}} - \mu}{\sigma/\sqrt{n}}$$

The Central Limit Theorem predicts that the distribution of normalised means should approximate the normal distribution.



The vertical lines show the centre points of the normalised mean and normal distributions.

As can be seen, the sample mean is very close to the theoretical mean, and the distribution of means is approximately normal, consistent with the Central Limit Theorem.