Manav Prabhakar
Shiv Nadar University



# Central Limit Theorem: Simulation, Visualization and Verification

Manav Prabhakar

December 16, 2020

## Problem Analysis

We need to take 'n' IIDs (independently and identically distributed random variables).

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Now, $\overline{X_n}$ is a statistic, or a function of the data. It is also a random variable, and the distribution associated with it is also referred to as the sampling distribution.

If $X_i$ belongs to normal distribution, with mean = 0, variance = 1, we need to plot it's pdf. We also need to simulate the sampling distribution for $\overline{X_n} = 1, 5, 25, 100$.

## Solution

Pdf for a normal distribution is given by

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Here,

$$\sigma^2 = 1, \mu = 0$$

We then get,

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Using this as the pdf, 1000 datapoints were taken and their pdf were calculated and plotted

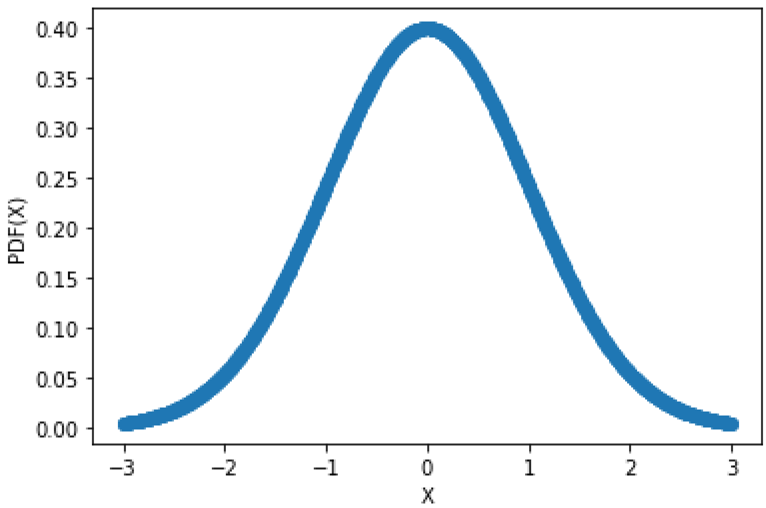


Figure 1: The figure shows a bell -shaped curve, a characteristic of normal distributions obtained using random number generator.

**Observations**

The sampling was done without replacement. We observe that as the sample size increases, the sample mean and sample variance converge to population mean and population variance.

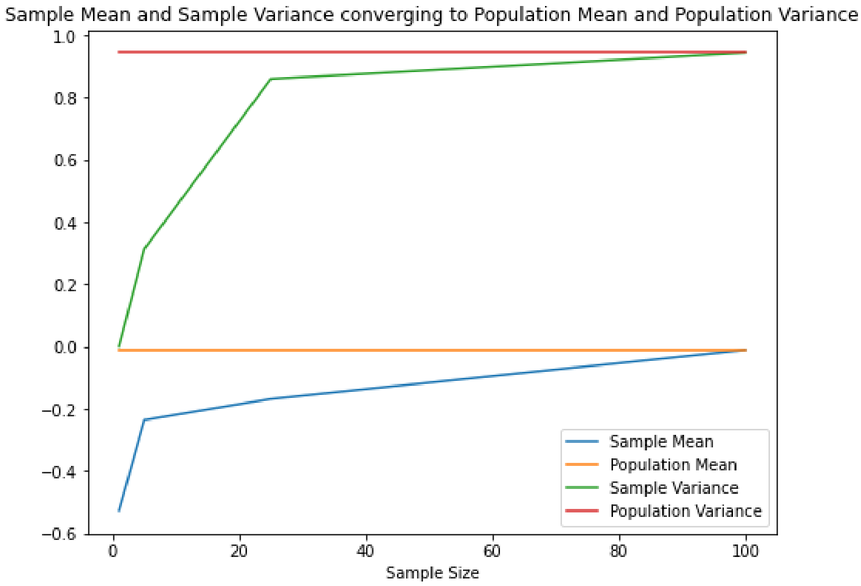


Figure 2: Sample mean and Sample variance converging to Population Mean and Population Variance

## Takeaways

The answers are similar to the theoretical answers. This problem is characteristic of the central limit theorem. The observations are randomly generated without affecting the generation of others, or we can simply say are independent. If we calculate the arithmetic mean for the observed values, then the probability distribution of this mean will be close to a normal distribution. It is because of this property that the means and variances are converging.

---------------------------------------------------x---------------------------------------------------