**The Central Limit Theorem**

Table of Contents

[Distribution of Sample Mean 3](#_Toc63699608)

In the Central Limit Theorem (CLT), we have IID random variables, and the sum of the random variables

We are interested in the distribution of . In reality, we will get an approximate distribution. gives us the sum of the values collected in a random sample. The CLT tells us that is approximately normal,

We need the mean,

Since the random variables have the same distribution function, .

Similarly, the variance is given by

Note that variances cannot be summed up like this generally. This is only applicable in this case since the random variables are independent.

Thus,

## Distribution of Sample Mean

Assume are IIDs and a random sample. Not all IIDs are random samples, but all random samples are IIDs. For this random sample, say the population mean is and the population variance is .

The sample mean is given by

is a function of random variables. As such, it is a random variable as well. This makes sense. Every time we collect a random sample, the value of the average data is going to change, which means the value of will also change.

Since is a random variable, we can have an expected value, variance, distribution etc.

Thus, the expected value of the sample mean is the same as the population mean.

The variance of the sample mean is the -th fraction of the population variance. This means that as we increase the value of , the graph for the variance of the sample mean will begin to converge. Eventually, this value will become when we take the data from the entire population.