

The Monte Carlo Simulation Algorithm for Sampling Distributions

1. Determine the population distribution of interest.
2. Use a pseudo random number generator (PRNG) to draw M samples each of size n from the population. (We are using the PRNG's builtin to Python's `np.random`)
3. Apply the sample statistic of interest to each of the M samples. Obtain M values of the sample statistic.
4. Plot a histogram of the simulated sample from the sampling distribution.
5. Calculate the mean and standard deviation of the simulated sample. If available (i.e., if the Central Limit Theorem, or CLT, applies compare with the theoretical values from the population.)

Glossary

The **Law of Large Numbers (LLN)** is a theorem that describes the result of performing the same experiment a large number of times. According to the LLN, the average of the results obtained from a large number of trials should be arbitrarily close to the expected value, and will tend to become closer as more trials are performed.

See the Wikipedia article on the Law of Large Numbers for more details.

The **Central Limit Theorem (CLT)** says that under random sampling, as the sample size gets large, the sampling distribution of the sample mean approaches a normal distribution with mean μ and variance σ^2/n . Put another way, if the sample size is sufficiently large, we can assume that the sample mean has a normal distribution. This means that with a 'sufficiently large' sample size, it can be assumed that

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

has a standard normal distribution.

Monte Carlo Methods or **Monte Carlo experiments** are a broad class of computational algorithms that rely on repeated random sampling to obtain numerical results.

See the Wikipedia article on Monte Carlo for more details.