Monte Carlo Method: Introduction

Junaid Hasan

Lecture 15: Discrete Math Modelling

Monte Carlo Methods

- The term **Monte Carlo** is used for a variety of numerical methods that involve randomness to solve problems.
- These methods are used for problems where a deterministic solution can be found, but it is either difficult to find or impossible to calculate.
- Therefore instead of finding the deterministic solution, a large sample is taken and an estimate is found.

Toy Example

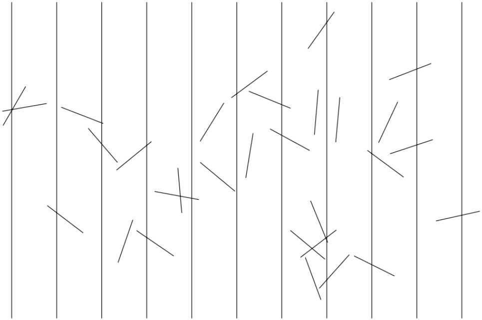
- Suppose you wanted to estimate the value of π.
- However, all you know that a circle has area πr^2 and you do not know anything else about π.
- A monte carlo method can be to inscribe a circle inside a square
- Then sample a large collection of points inside the square uniformly.
- The ratio of points inside the circle vs the entire square gives an estimate for $\pi/4$.
- Let us try this out in Python

General Overview of MCM

- Define a domain of possible inputs (circle inside the square).
- Generate inputs randomly from a probability distribution over the domain.
- Perform a deterministic computation (taking ratio of points in vs out) on the inputs.
- Aggregate the results.

Buffon's Needle

- Suppose we have a surface lined with equally-spaced parallel lines.
- We have a needle whose length equals the space between parallel lines.
- If we throw the needle, what is the probability it will cross a line.
- For example something like this could happen



-

Applying MCM

- We must decide on a domain to work on.
- We will probably sample uniformly on the domain.
- Then we will use the ratio number of needles cross/ number of needles thrown
- To estimate.
- The more needles we throw the better estimate.

Deciding the domain

- Deciding the domain is tricky.
- Any ideas?

A possible solution

- Say we assume that **one end of the needle will land uniformly between two lines**.
- Furthermore we assume space to be $1 =$ length of the needle.
- This means that one end of the needle will land uniformly in $[0, 1]$.
- The other assumption is that the angle made by the needle with the one end as the center is uniform over $[0, 2\pi)$.
- Therefore we can estimate angle by another random number generator between $[0, 2\pi)$.

Approximate pi

- The Buffon needle simulation should give the fraction as a close approximation of $\frac{2}{\pi}$.
- Furthermore, this can help estimate π.
- The catch is that often times we will need multiple runs (say 10000, or million) to get good results.
- We can try plotting

Benefits

- MCM are often used to estimate probabilities that might otherwise be difficult to work out with exact methods.
- For example, if we want to estimate the probability of winning a game of solitaire using clearly-defined rules and strategies it may be quite challenging if we want an exact value.
- With solitaire, and a fixed strategy, all depends on the permutation of the deck. Since there are $52! \approx 10^{67}$ permutations, it is not possible to try all possible decks and count which leads to win/loss.

contd..

- Instead we can use a simulation to generate random permutations of the deck, and then play the game with the strategy we are investigating to see if each permutation leads to a win or not.
- The ration number of wins/number of games played will converge to the probability of a win as the number of games tend to infinity.
- This is because of the Law of Large Numbers.
- We can run the simulation as many times as we need until the estimate gives us some confidence in its being close to the actual probability of winning with this strategy.

contd..

- If there are multiple strategies that can be applied to a game, simulation is often a great way to compare strategy and compare win probabilities.
- Deciding how many runs are needed is a bit of an art, but often estimates can be found using Central Limit theorems.

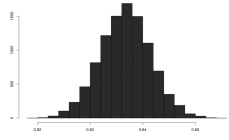
Buffons Needle and Central Limit Theorem

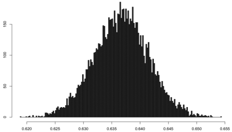
- An important theorem in probability is the Central Limit Theorem. It says (roughly) that if we sample many values from an unknown distribution and take their mean, and do this multiple times, the means will be approximately **normally distributed**.
- A normal distribution (also known as **Gaussian**) has the shape of $y = e^{-x^2}$ with scaling.
- Features: It is symmetric about its peak, it has exactly two inflection points, and is asymptotic to the x-axis.

contd..

- Suppose we throw the needle in Buffon's needle 10000 times.
- We can view this as sampling from \${0,1 } bernoulli trials with 1 meaning the needle crosses the line and 0 meaning it does not.
- Thus throwing 10000 times and calculating empirical probability can be seen as averaging 10000 random samples from $\{0, 1\}$ with the probilities given.
- If we do this many times, and make a histogram of the results we will notice a normal shape.

Experiment

- Running multiple runs of 10000 needle throws and recording the estimated probability from each run yeilds something like
- 

- 

CLT lessons

- The CLT says what these histograms suggest that the distribution of these values is very nearly normal.
- Furthermore we can use statistics now to get confidence intervals.
- We will do this next time.