Monte Carlo Sampling

Shorya Kumar (190818), Saaransh Aggarwal (190729)

1 April 2022

Example

Lets start with an example where we can use Monte Carlo approximation , lets try to estimate the value of π using Monte Carlo Simulations. To begin with lets consider a square of unit radius circumscribing a quarter circle. We will use this set up to estimate the value of π. As we know how to sample from a uniform distribution we sample both X and Y from uniform(0,1) i.e. $X \sim U(0,1)$ and $Y \sim U(0,1)$. Now this point will lie inside the quarter circle with probability $\frac{(area of quarter circle)}{(area of square)}$, which is equal to $\frac{\frac{\pi}{4}}{1} = \frac{\pi}{4}$. Now we will exploit this to estimate π. We can define an Indicator Random Variable Z as:

$$Z = 1 \ : \ (X^2 + Y^2 \leq 1)$$

$$Z = 0 \ : \ (X^2 + Y^2 > 1)$$

If we try to calculate Expected Value of Z such that $X \sim U(0,1)$ and $Y \sim U(0,1)$, it comes out as

$$E_{X \sim U(0,1) \ \ Y \sim U(0,1)}[Z] = \int\limits_{X^2+Y^2 \leq 1} Z * dXdY + \int\limits_{X^2+Y^2 < 1} Z * dXdY$$

$$= \int\limits_{X^2+Y^2 \leq 1} 1 * dXdY + \int\limits_{X^2+Y^2 < 1} 0 * dXdY = \int\limits_{X^2+Y^2 \leq 1} dXdY = \frac{\pi}{4}$$

Now we can use Monte-carlo to calculate the expected Value of Z and compare it with $\frac{\pi}{4}$ to estimate π. Lets now get into what is Monte carlo.

1 Generalization

$$E_{x \sim P_X(x)}[f(x)] \ \approx \ \frac{\sum_{s=1}^{M} f(x_s)}{M} \ \ ; \ x_s \sim P_X(x)$$

So crudely what Monte Carlo approximation states is that expected value of a function f(x) when x is sampled from a distribution p(X) is equal to the average of f(x) computed on samples from p(X).

$$E_{x \sim P_X(x)}[f(x)] \ \approx \ \sum_{x \in S} p_X(x) f(x) \ \approx \ \frac{\sum_{s=1}^{M} f(x_s)}{M} \ \ ; \ x_s \sim P_X(x)$$

eg: for $S = H, T$

$$p(H)f(H) \ + \ p(T)f(T) \ \approx \ \frac{\sum_{s=1}^{M} f(x_s)}{M} \ \ ; \ x_s \sim P_X(x)$$

Some uses of MC approximations are listed below:

1. Bayesian Inference (details left)
2. M step of EM algorithm (details left)

2 Sampling

Given a $U_X(x)$ uniform random number generator , how to sample from other distributions.
Special case N(x,0,1) To sample from $N(x,0,1)$ i.e standard normal distribution, we know that by CLT sum of iid random variables converge to gaussian.
If x_i is a distribution with mean μ_i and variance σ_i and all $x_i s$ are independent and identically distributed:

$$\sum_{i=1}^{n} \frac{x_i - \mu_i}{\sigma_i} \sim N(x,0,1)$$

Now lets look at some other ways to sample

2.1 Markov Chain Monte carlo

Let us say we have 2 states s_1 and s_2 and we are given probabilities of transition from one state to the another and also the emission probabilities of remaining in the same state. Our aim is to calculate the equilibrium state. There are 3 ways to solve the given problem.

- Using Conditional Probability

  $p(x^{(3)} = s_1) = \sum_{x^{(2)}} p(x^{(3)}|x^{(2)})p(x^{(2)})$

  $p(x^{(3)} = s_2) = \sum_{x^{(2)}} p(x^{(3)}|x^{(2)})p(x^{(2)})$

  Similarly, we can keep on calculating $x^{(4)}, x^{(5)}, ...$ and get the equilibrium value.

- Using Monte Carlo Sampling
  1. Start from a random state
  2. Perform Monte Carlo Sampling multiple times from $p(x^{(t+1)}|x^{(t)})$
  3. After 1000 iterations, note down the state.
  4. Repeat the steps 1-3 multiple times, and the statistics will give the value of $p(x^{(eq)} = s_1)$ and $p(x^{(eq)} = s_2)$

- Analytical Calculation

  Let us assume that the equilibrium distribution is $p(x)$.

  Now, $p(x^{(t+1)} = s_1) = p(x^{(t)} = s1)p(x^{(t+1)} = s_1|x^{(t)} = s_1) + p(x^{(t)} = s_2)p(x^{(t+1)} = s_1|x^{(t)} = s_2)$

  At the equilibrium state, we assume $p(x^{(t+1)}) = p(x^{(t)})$

  Substituting $p(x^{(t+1)}) = p(x^{(t)})$ in the above equation, we can find the values of $p(x^{(eq)} = s_1)$ and $p(x^{(eq)} = s_2)$ directly.

2.1.1 Convergence of Markov Chain

Markov Chain will converge to a distribution $p(x)$ if the following condition holds

$p(x^{'}) = \sum_x T(x- > x^{'})p(x)$

The advantage of using Monte Carlo Markov Chain is that if the given $p(x)$ is difficult to sample from, we can use an appropriate transition probability to sample from the given distribution.

2.1.2 How to Construct Markov chains

Gibbs Sampling :

1. Start with any $x_1^{(0)}, x_2^{(0)}, x_3^{(0)}$
2. Sample as follows:
   $x_1^{'} \sim p(x_1|x_2 = x_2^{(0)}, x_3 = x_3^{(0)})$
   $x_2^{'} \sim p(x_2|x_1 = x_1^{(1)}, x_3 = x_3^{(0)})$
   $x_3^{'} \sim p(x_3|x_1 = x_1^{(1)}, x_2 = x_2^{(1)})$
3. Repeat 2

Reference:

- EE698R Lectures