COMP 480/580 — Probabilistic Algorithms and Data Structure Oct 20, 2022

Lecture 17

Lecturer: Anshumali Shrivastava *Scribe By: Wenze Ma, Jiehui Li and Xilin Song*

1 Review: Monte Carlo Simulation

1.1 Definition

Monte Carlo Simulation is a method used for the prediction between multiple results under the situation that evolves randomly. Basically, a sample is randomly chosen, and the sample will fall in one of the groups. With more samples chosen, the proportion between the groups will be more closed to the expected value.

1.2 Example

We can use Monte Carlo's method to estimate the value of π. For example, we have a square dartboard with length of $2r$. There is also a circular area that fits perfectly inside the square with a radius of r. Then we begin throwing darts to the board repeatedly. And we will keep track of the number of darts landing in the square and the circle, as demonstrated in Figure 1.

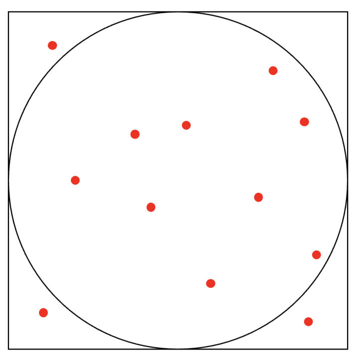

Figure 1: Monte Carlo dart throwing example.

Let the total number of darts be A, and the number of darts landing inside the circle be a. Then, we can estimate π such that $\frac{a}{A} = \frac{\pi}{4}$
In general,

$$\frac{a}{A} = \mu$$

where μ is the expected value.

1.3 Computational Cost

The only cost in the simulation is to generate samples. The number of samples will also affect the accuracy of the prediction. So, we need to consider the error ϵ when determine the number of samples we need. When we consider the error, we can find the range of the estimated expected value.

$$(1-\epsilon)\mu \leq \hat{\mu} \leq (1+\epsilon)\mu$$

The overall computational cost will be equal to the cost of generating N samples with an error of ϵ. So the overall cost will be

$$N \geq \frac{1}{\mu} O(\frac{1}{\epsilon^2})$$

2 DNF Counting

2.1 Definition

DNF is a conjunction of clauses, where each clause is a disjunction of x_i, where x_i is a Boolean variables. For example, $(x_1 \wedge x_5 \wedge \overline{x_9}) \vee (x_4 \wedge x_5 \wedge x_6) \vee ...$ is a DNF.

DNF counting is a problem counting the number of satisfying assignments for variables. Since the clauses are connected by the logical OR, the whole statement will be true if there is at least one satisfying clause. The statement will not be satisfied if and only if all clauses are not satisfied.

2.2 DNF Properties

For DNF

$$\underset{C_1}{(x_1 \wedge x_2 \wedge x_3)} \vee \underset{C_2}{(x_4 \wedge x_5 \wedge x_6)} \vee ... \vee \underset{C_M}{(x_{n-2} \wedge x_{n-1} \wedge x_n)}$$

Let S_{C_i} = set of the assignment satisfying clause C_i. Then,

$$|S_{C_i}| = 2^{n-k_i}$$

where k_i is the number of satisfying clause

Let S be the set of satisfying assignment such that S is the union of S_{C_i} for $1 \leq i \leq M$. Then we can say $(i, S_{C_i}) = \{(i, A) : A \in S_{C_i}\}$ and

$$\bigcup_{i=1}^{M} (i, S_{C_i}) = S^*$$

Since $|S^*| \leq M \cdot |S|$, we can conclude that S^* is a set that is not too big.

For the set $s^* = \{(i, a) : a \in S_{C_i} \forall k < i, a \notin S_{C_k}\}$,

$$|s^*| = |S|$$

2.3 How to uniformly sample from S^*?

We have two sets S_1 and S_2 with size of n_1 and n_2 respectively, where $S_1 \cap S_2 = \emptyset$. To uniformly sample from S_1 and S_2, the estimation will be

$$\frac{n_1}{n_1 + n_2}$$

3 Markov Chain, Stochastic Process

3.1 Definition

Stochastic process is a sequence of states. Markov chain is a sequence of random variables, which take value in the in the states $\{X_0, X_1, X_2, ..., X_t, ...\}$

For example,

$$\Omega = \begin{cases} 1, 2, 3, ..., N \\ 2, 1, 3, ..., N \\ \vdots \\ N, N-1, ..., 1 \end{cases} \tag{1}$$

Ω is the set of all states and there are $N!$ states in total.

3.2 Markov Chain Property

The property of Markov chain is memory-less, that the next state is only dependent on the current state, and past states does no matter.

$$Pr(x_t = y_t | x_{t-1} = y_{t-1}, x_{t-2} = y_{t-2}, ..., x_0 = y_0) = Pr(x_t = y_t | x_{t-1} = y_{t-1}) = Pr(y|x_t)$$

3.3 State Transition Graph

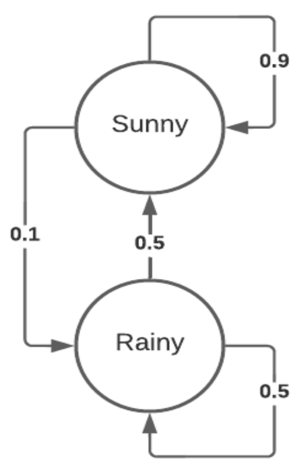

Figure 2: State transition graph example

The graph is a representation of Markov chain since today's weather completely depends on yesterday's weather. And yesterday's weather was only dependent on the weather of the day before yesterday and so on. For example, if yesterday's weather is sunny, then there is a 0.9 probability that today is also sunny, and a 0.1 probability that today is rainy. Notice that these two numbers add up to 1.

$$Pr(Sunny|Sunny) = 0.9 \qquad Pr(Rainy|Sunny) = 0.1$$