Introduction

Bo Li, Spring 2019

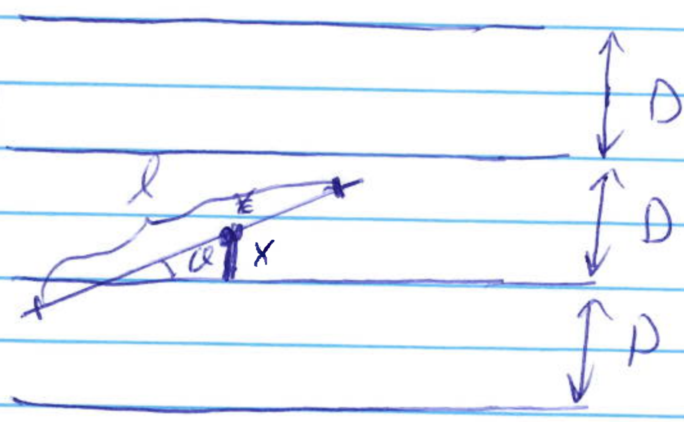
Monte Carlo Methods are a class of computer simulation methods using random numbers. They are used for simulating directly systems that are naturally random or fluctuating. They are also used to simulate artificial random systems to determine or approximate deterministic quantities.

Monte Carlo methods have been used extensively in physics, chemistry, and biology, and in statistics, operations research, and finance.

(A) Examples

Example 1 (Buffon's needle problem, 1977) This is an example of Monte Carlo (non-computer) experiment, used to determine a deterministic quantity.

Given parallel lines (with separation D) on ground. Throw needles on the ground.



x = dist. of center of needle to lines

l = length of needle

θ = acute angle between needle and line.

$(l < D)$

X, θ are both random variables.

X is distributed uniformly. The probability distribution of X is $2/D$ if $0 \le x \le D/2$, 0 otherwise. θ is uniformly distributed, with the distribution being $2/\pi$ if $0 \le \theta \le \frac{\pi}{2}$, 0 otherwise. The joint distribution of (X, θ) is uniform on $0 \le x \le \pi/2$, $0 \le \theta \le \frac{\pi}{2}$, i.e., the joint distribution is given by $4/D\pi$ if $0 \le x \le \pi/2$, $0 \le \theta \le \frac{\pi}{2}$ and 0 otherwise. The needle crosses a line if $x \le \frac{l}{2} \sin\theta$.

The probability that a needle crosses a line is then

$$p = \int_0^{\frac{\pi}{2}} \int_0^{\frac{l}{2}\sin\theta} \frac{4}{D\pi}\, dx\, d\theta = \frac{2l}{\pi D}.$$

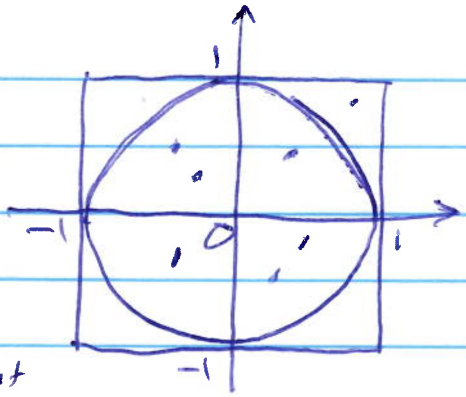
Hence, $\pi = \frac{2l}{pD}$. On the other hand,

$$p \approx \frac{\#\text{ needles crossing lines}}{\#\text{ needle}}$$

if the # needle $>> 1$. More precisely, let P_N denote the proportion of "intersects" in N throws. We have

$$\pi = \lim_{n \to \infty} \frac{2l}{P_N D}.$$

Example 2 Computing π again, using a Monte Carlo computer simulation this time.



Area of circle $= \pi$
Area of square $= 4$.
Randomly throw a point onto the square. The probability that this point falling into the circle is $\pi/4$.

Suppose we repeat this "experiment" $N (\gg 1)$ times. Let I_N be the number of times the point is inside the circle. Then, natrally $I_N/N \approx \frac{\pi}{4}$. i.e., $\lim_{N\to\infty} 4\, I_N/N = \pi$. Indeed, these N points are uniformly and independently distributed on the square. The random variable I_N has a binomial distribution with the parameters N and $\pi/4$. Therefore

$$E(4 I_N/N) = \pi.$$

This means $4 I_N/N$ is an unbiased estimator.

This method can be generalied to the calculation of area, volume, etc. using Monte Carlo methods.

Continuing, we can look into the accuracy issue. This is related to the efficiency issue. So, we ask, if we want to have the estimate to have an error within certein range, how large N we need.

Let X_k be the random variable: $X_k = 4$ if the point in the kth "throwing experiment" lies inside the circle, and 0 otherwise. Then $X_1, X_2, \cdots, X_n, \cdots$ are independent and identically distributed (i.i.d.) with the expectation $\mathbb{E}(X_k) = \pi$ $(\forall k \geq 1)$. Define

$$\overline{X}_N = \frac{1}{N}(X_1 + \cdots + X_N).$$

call it a sample mean. Then the Law of Large Numers (LLN) implies that $\overline{X}_N \to \mathbb{E}(X_1) = \pi$ with probability 1. Moreover, the Central Limit Theorem (CLT) implies that the sequence

$$(\overline{X}_N - \pi)/(\sigma/\sqrt{N})$$

converges to a random variable that has the standard normal distribution. Here $\sigma^2 = \mathrm{Var}(X_k)$ $(\forall k \geq 1)$ is the variance (equivalently, σ is the standard deviation). Given $\varepsilon > 0$, an error torelance, we have then for $N \gg 1$ that the probability

$$\mathbb{P}\left(-\varepsilon \leq \frac{\overline{X}_N - \pi}{\sigma/\sqrt{N}} \leq \varepsilon\right) \approx \int_{-\varepsilon}^{\varepsilon} g(x)\,dx$$

where $g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ $(x \in \mathbb{R})$ is the probability density function (PDF) of the standard Gaussian (i.e., normal) rand variable. We can rewrite this as

$$\mathbb{P}\left(|\bar{X}_N - \pi| \le \frac{\varepsilon\sigma}{\sqrt{N}}\right) \approx \int_{-\varepsilon}^{\varepsilon} g(x)\,dx.$$

Thus, for $N >> 1$, the probability that π (which is unknown in our "experiment" and analysis) falls into the interval

$$\left[\bar{X}_N - \frac{\varepsilon\sigma}{\sqrt{N}},\ \bar{X}_N + \frac{\varepsilon\sigma}{\sqrt{N}}\right]$$

is approximately

$$\gamma := \int_{-\varepsilon}^{\varepsilon} g(x)\,dx \in (0,1).$$

This γ-value (percentage) can be numerically estimated for a given $\varepsilon > 0$. We call the interval *γ-confidence interval* for the estimator $\bar{X}_N$ of π.

Note that, in this case one has $\sigma^2 = \pi(1-\pi)$. But, π is supposed to be unknown in this analysis. So, we need estimate it. We can estimate $\sqrt{x(4-x)}$ for $x \in (0,4)$, e.g., $\sigma \le \sqrt{4 \cdot 4} = 4$.

Example 3 Let D be a bounded region in $\mathbb{R}^d$ and $g : D \to \mathbb{R}$ a continuous and bounded function. We consider the integral

$$I = \int_D g(x)\,dx.$$

If $d >> 1$ then the usual numerical integration methods will be prohibited. We should try to use Monte Carlo methods. The starting point of Monte Carlo integration is to reformulate the integral as the expectation of some random variable. Let $f : D \to \mathbb{R}$ be such that $f \geq 0$ on D. ~~(but $f \not\equiv 0$)~~ and $\int_D f(x)dx = 1$. So, f is the probability density function of some random variable $Z \in D$. A simple choice of f is $f(x) = 1/|D|$ $(x \in D)$, where $|D|$ is the volume of D (if that is known or easy to calculate). Now, we have

$$I = \int_D \frac{g(x)}{f(x)} f(x) dx = \int_D h(x) f(x) dx,$$

$h(x) = g(x)/f(x)$. Hence

$$I = \mathbb{E}_f(h(\cancel{X}Z)),$$

i.e., the expectation of random variable $h(Z)$ with respect to the density f.

The next step is to generate random variables $X_1, X_2, \cdots, X_k \cdots$:

(1) All $X_k \in D$.

(2) $X_1, \cdots, X_N, \cdots$ are independent.

(3) Each X_k is distri~~buted~~buted according to f.

Now, by the Law of Large Numbers (LLN)

$$I = \mathbb{E}(h(Z)) = \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} h(X_k).$$

So, for $N >> 1$, $I \approx \frac{1}{N}\sum_{k=1}^{N} h(X_k)$.

The error in this estimation is

$$\text{error} = \left| \frac{1}{N}\sum_{k=1}^{N} h(X_k) - I \right|$$

$$= \frac{\sigma_h}{\sqrt{N}} \left| \frac{\frac{1}{N}\sum_{k=1}^{N} h(X_k) - I}{\sigma_h / \sqrt{N}} \right|,$$

where $\sigma_h^2 = Var(h(X_k))\ (\forall k \geq 1)$. We have

$$\sigma_h^2 = Var(h(Z))$$

$$= \int_D [h(x) - I]^2 f(x) dx$$

is the variance of $h(Z)$, and hence $h(X_k)$ for all $k \geq 1$. The Central Limit Theorem (CLT) implies that, for $N >> 1$,

$$\frac{\frac{1}{N}\sum_{k=1}^{N} h(X_k) - I}{\sigma_h / \sqrt{N}}$$

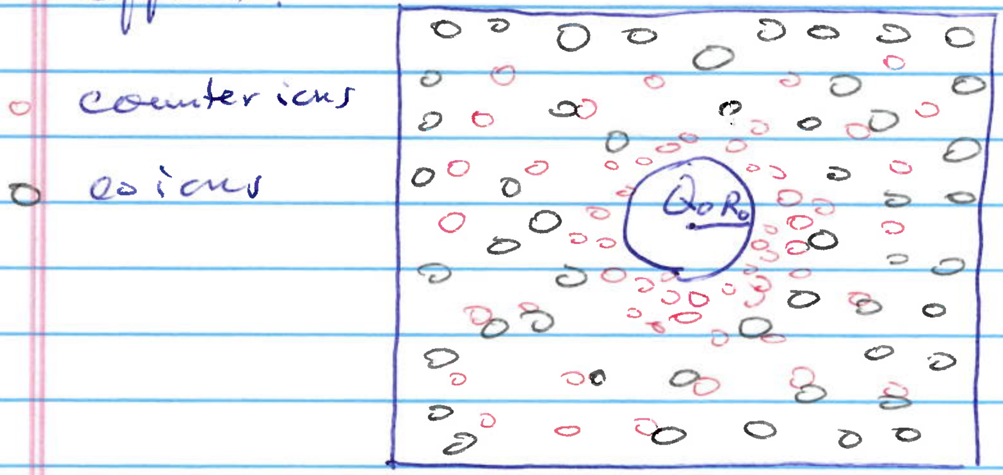
is distributed approximately according to the standard Gaussian $\mathcal{N}(0,1)$ As before, we can determine a γ-confindence interval for estimating I by $\frac{1}{N}\sum_{k=1}^{N} h(X_k)$. The question is again: how to estimate the variance σ_g^2 as that is unknown.

Monte Carlo integration is important as integrals represent the means of random variables. From this example we see immediately there are two questions.

(1) How to generate random variables?

(2) How to reduce the variance? (see, e.g., the term σ_h in the error expression on page $\boxed{7}$).

<u>Example 4</u> Monte Carlo simulations of ionic systems to investigate the ionic size effect.



Consider the region $(-L,L)^3$, our simulation box. At the center (0,0), we place (and fix it) a relatively large colloidal ball of radius R_0 (e.g., $R_0 = 20$ Å), and also place a charge Q_0 (>0) at the center (0,0). This is an effective point charge of the central ball. (The point charge can be equivalently redistributed as the surface charges on the surface of the ball.)

We put many mobile ions in the box, say, M types (or species) of them.

Each ion of type i carries a point charge $Q_j^{(i)}$ and has the radius $R_j^{(i)}$, $j=1,\cdots,N_i$, $i=1,\cdots,M$, where N_i is the total number of ions of type i.

The solvent (e.g., water) is treated implicitly through the dielectric coefficient $\varepsilon > 0$.

We require the charge neutrality

$$Q_0 + \sum_{i=1}^{M}\sum_{j=1}^{N_i} Q_j^{(i)} = 0 .$$

Often $M = 2$ or 3 or 4, or 5. $Q_j^{(i)}$ can be $+1, +2$, and $-1, -2, -3$. Moreover, N_i can be a few thousands to more.

Let us relabel all the mobile ions as $1, \cdots, N$. So, their positions, radii, and charges are labeled by

$$\vec{r}_k = (x_k, y_k, z_k),\ R_k, \text{ and } Q_k \quad (1 \le k \le N).$$

We set $\vec{r}_0 = (0,0,0)$.

~~[We use Monte Carlo methods to simulate this system to get equilibrium distributions.]~~

All the ions (including the central ball) interact through a potential $U = U(\vec{r}_0, \vec{r}_1, \dots, \vec{r}_N)$, given by

$$A\,U(\vec{r}_0, \vec{r}_1, \dots, \vec{r}_N) = \sum_{i,j=0,\, i\neq j}^{N} U_{ij}(|\vec{r}_i - \vec{r}_j|),$$

where A is a constant that depends on the temperature T, dielectric coefficient ε, etc.,

$$U_{ij}(|\vec{r}_i - \vec{r}_j|) = \begin{cases} \dfrac{Q_i Q_j}{|\vec{r}_i - \vec{r}_j|} & \text{if the balls } B(\vec{r}_i, R_i),\ B(\vec{r}_j, R_j) \text{ do not intersect,} \\ +\infty & \text{otherwise.} \end{cases}$$

We use the periodical boundary condition.

We use the Metropolis algorithm to simulate this system to get the equilibrium distributions of different types of mobile ions. (Intuitively, counterions will be dense around the central ball, and coions will be mostly away from the central charge.) This algorithm is perhaps the first modern Monte Carlo algorithm, and is also perhaps the most popular Monte Carlo method.

The method generates a sequence of configurations $X_k = (\vec{r}_0, \vec{r}_1^{(k)}, \dots, \vec{r}_N^{(k)})$ (note that $\vec{r}_0 = (0,0,0)$ is fixed). The sample means

$$\overline{X_N} = \frac{1}{N}\sum_{k=1}^{N} X_k$$

provide good estimates of the equilibrium configuration as $N >> 1$.

The algorithm

Given $X^{(k)} = (\vec{r}_0, \vec{r}_1^{(k)}, \cdots, \vec{r}_N^{(k)})$.

(1) choose $J \in \{1, \cdots, N\}$ uniformly at random.

(2) Define $Y = (\vec{r}_0, \vec{r}_1, \cdots, \vec{r}_{J-1}, \vec{r}_J^{\,*}, \vec{r}_{J+1}, \cdots, \vec{r}_N)$ by $\vec{r}_J^{\,*} = \vec{r}_J^{\,(k)} + \vec{\Delta r}$ where $\vec{\Delta r} = (\Delta x, \Delta y, \Delta z)$ with $\Delta x, \Delta y, \Delta z$ independent and uniform on $(-l_x, l_x)$, $(-l_y, l_y)$, $(-l_z, l_z)$, respectively. Here, l_x, l_y and l_z are prefixed maximal pertarbation sizes (or more sizes).

(3) Define $\alpha = \cancel{\alpha} \min\left(1, e^{-[U(Y) - U(X^{(k)})]}\right) \in (0,1)$

And accept Y with probability α. i.e.,

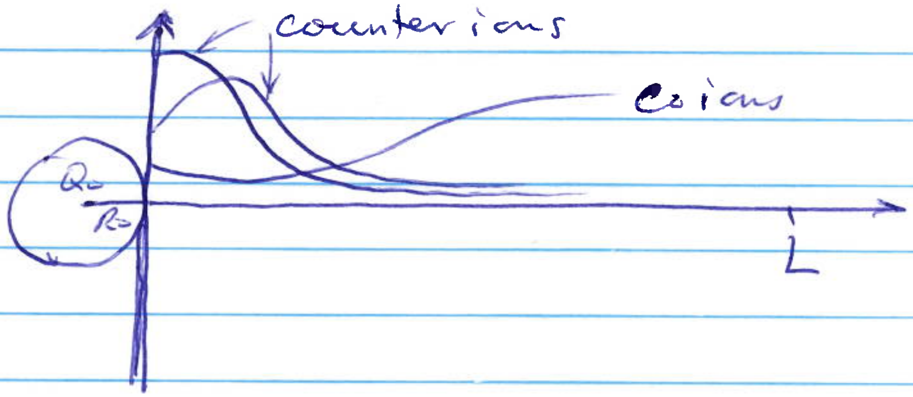
Generate a random number U that is uniformly distributed on $[0,1]$.

If $U \le \alpha$ then accept Y: $X^{(k+1)} \longleftarrow Y$.

If $U > \alpha$ then reject Y. and $X^{(k+1)} \longleftarrow X^{(k)}$.

The sequence $\{X^{(k)}\}_{k=0}^{\infty}$ is a Markov chain. It is reversible, i.e., satisfying the detailed balance. Moreover, it is ergodic so that each configuration will be visited infinitely many times (in principle).

Once the Markov chain is obtained, we need to process~~ing~~ the simulation data. For instance we can plot histogram to show the particle density for each type of mobile ions. ~~For~~ A typical result (after smoothing out the histogram) is a plot of densities



(B) Basic Questions / Topics

Questions

(1) Design estimators for some quantities
Algorithms.

(2) Accuracy and efficiency.

(3) Data processing.

Topics

⊙ Generating (or sampling) random variables according to a give distribution, continuous or discrete. This is a basic step in a Monte Carlo method. Here, one generates a sequence of random variables, e.g., i.i.d. sequence or a Markov chain, according to distribution. There are many methods for generating a random variable. The starting point is to use existing uniform random number generators. [In these notes, we do not discuss how uniform random numbers can be generated — on computers.]

⊙ Variance reduction. While there are many methods to sample random variables. Some of them may give large variances and others small variances. Techniques that can reduce the variance will then be useful.

⊙ Markov chain Monte Carlo. A class of methods for generating Markov chains with given targeted distributions. Some of the well-known methods include: The Metropolis method, or a generalized version, the Metropolis-Hasting Method, the Gibbs sampler, the simulated annealing method, etc.

⊙ Statistical analysis of simulation data, including methods/techniques of presenting data, postprocessing data to get better results, etc.

Additional topics include: Monte Carlo integration and summation — which are largely in the part variance reduction; Monte Carlo optimization; rare event simulation; and application in statistical mechanics.

I have collected some basics of probability and Markov chains in the Appendix. Useful analysis tools include: LLN, CLT, and other random variable convergence techniques. Martingales will be also useful sometimes.

It is always good to look into carefully some useful examples/systems to under the method. These include: A particle moves in a potential; a many-particle system; random walk on a graph; some combinatorial optimization problems (such as the traveling salesman problem, the knapsack problem, etc.

⊙ N. Madras, Lectures on Monte Carlo Methods, Amer. Math. Soc., 2002
[A well written, concise introduction and only 103 pages!]

⊙ C. P. Robert and G. Casella, Monte Carlo Statistical Methods, 2nd ed., Springer, 2004
[Quite comprehensive. More for statisticians]

⊙ R. Y. Rubistein and D. K. Kroese, Simulation and the Monte Carlo Method, 3rd ed., Wiley, 2017.
[Quite comprehensive, many examples, algorithms. More for statisticians. With exercise problems.]

⊙ P. J. M van Laarhoven and E. H. L. Aarts, Simulated Annealing, Springer, 1989.
[Specialized. Convergence analysis.]

Application in Physics and Chemistry

⊙ M. P. Allen and D. J. Tildesley, Computer Simulation of Liquids, Oxford University, 1987.

⊙ K. Binder and D. W. Heermann, Monte Carlo Simulation in Statistical Physics, 5th ed., Springer, 2010.

References

General

- S. Asmussen & P.W. Glynn, Stochastic Simulation, Springer, 2007.
  [Algorithms and analysis. covers a lot of topics. With exercise problems.]
- G.S. Fishman, Monte Carlo: Concepts. Algorithms and Applications. Springer, 1996
  [An earlier reference. Algorithm and analysis.]
- M.H. Kalos and P.A. Whitlock, Monte Carlo Methods, ~~2nd ed~~ 2nd ed. Wiley, 2008.
  [A nice introduction to Monte Carlo Methods, with application in statistical mechanics]
- D. P. Kroese, T. Taimre, and Z.I. Botev, Handbook of Monte Carlo Methods. Wiley, 2011.
  [It is a HANDBOOK, with computer codes. Good as a reference book.]
- J.S. Liu, Monte Carlo Strategies in Scientific Computing. Spring, 2008.
  [For statisticians and computational scientists. Some applications to biomolecular systems, With exercise problems.]

- D. Frenkel and B. Smit, Understanding Molecular Simulation, 2nd ed., Academic Press, 2002
- D. P. Landau and K. Binder, A Guide to Monte Carlo Simulations in Statistical Physics, 3rd ed., Cambridge University Press, 2009.

Mathematical Foundation

- C. Graham and D. Talay, Stochastic Simulation and Monte Carlo Methods, Springer, 2013.