Chapter 1

OVERVIEW

9/13/2013

1.1 INTRODUCTION

The "Monte Carlo Method," also known as the "Method of Statistical Trials," is a statistical method for solving deterministic or probabilistic problems across the scientific and engineering disciplines.

It is generally applied in two steps. First, the solution of a problem is represented as a parameter of a hypothetical population. Second, a random sequence of numbers is used to construct a sample of the population, from which statistical estimates of the needed parameters can be obtained.

1.2. GENERAL DESCRIPTION OF THE MONTE CARLO METHOD

Monte Carlo can be considered as a method of solving various problems in computational mathematics by constructing for each problem a random process with parameters equal to the required quantities of the initial problem. The unknowns are determined by carrying out observations on the constructed random process, and by computing its statistical characteristics, which would be equal to the required parameters of the initial process.

More simply stated, it is a numerical method for solving mathematical problems by means of random sampling.

In even simpler terms, it is a physics experiment carried out numerically on a computing platform, rather than in the laboratory or the real world.

Several excellent publications have been published about the method. These include a monograph by Hammersley and Handscomb, a report by Kahn, a review by Halton, and the books by Shreider, Sobol, Spanier and Gelbard, and by Carter and Cashwell.

1.3. HISTORICAL ORIGINS

The Monte Carlo method can be traced back to Pharaoh, Babylonian and Old Testament times. Mention should be made of early ideas due to Laplace in 1812, Lord Raleigh in 1899, Student in 1908, Courant, Friedrichs, and Lewy in 1928, Kolmogorov in 1931, and Pólya in 1938.

The method has its basis in probability theory, whose larger strides were started in the 16^{th} century. Engineers and scientists have discovered that the laws governing randomness can describe the real world. Probability theory dates back to Girolamo Cardano(s) (1501-1576), who wrote in 1526 "The Book of Games of Chance," or "Liber de Ludo Aleae." Cardano, famous for his cubic equation, was a mathematician, engineer,

physician, as well as a gambler. Probability theory was developed by Jacques Bernoulli I (1654-1705), Pascal, De Moivre, Euler, Laplace, Gauss, and Poisson (1781-1840). A large number of mathematicians developed it further in the 19th and 20th centuries.

Lord Raleigh in 1899 showed that a one dimensional random walk without an absorbing barrier can provide a solution to a parabolic differential equation. A. N. Kolmogorov (1903-1987) showed in 1931 the relationship between Markov chains stochastic processes and the solution to some integro-differential equations. William Sealy Gosset (1876-1937), better known as Student, in 1908 used experimental sampling to help him in the discovery of the distribution of the correlation coefficient, as well as checking his earlier derivation of his famous Student t-distribution.

The systematic use of the method, and its name, date back to the 1940s period of the Los Alamos School of mathematicians and physicists, and especially to the work of von Neumann (1903-1957), Stanislav Ulam, Metropolis, Enrico Fermi, Kahn, and Richtmyer.

The expression "Monte Carlo Method," was first used in 1946 at the Los Alamos National Laboratory in New Mexico, USA. It was applied to a computational program developed by John von Neumann along lines suggested by Stanislav Ulam, in which a neutron chain reaction is simulated on a digital computer, with the aid of random numbers, and the simulated reaction then studied by standard statistical methods. At this time, John von Neumann and Stanislav Ulam introduced variance reduction methods such as the "Splitting Method" and "Russian Roulette".

The problem during World War II pertained to how far would neutrons travel in different materials, and had a vital bearing on shielding and the calculation of critical masses. To explore the problem by experimental trial and error would have been quite expensive, time consuming and outright hazardous. The problem also seemed beyond the reach of theoretical calculations. The physicists knew, however most of the basic data, such as the average distances or mean free paths that a neutron of a given speed or energy would travel in a given material before it collided with an atomic nucleus. They also knew the probabilities of neutron scattering, absorption, and fission, as well as the energy loss of the neutrons after each collision. It was impossible to sum up this information into a practical formula for predicting the outcome of a whole sequence of neutron interaction events. Stanislav Ulam and John von Neumann approached the problem by sampling the probabilities of the process step by step. The separate events were merged into a composite picture, which gave a workable answer to the problem. This mathematical technique was known to statisticians for years. When it was revived for the work at Los Alamos, John von Neumann gave it the war-time code name of "Monte Carlo."

After the war the same type of computations were used in the "Super" program, or the construction of hydrogen or thermonuclear devices. The calculations were carried out by John von Neumann, at Princeton University, on the MANIAC computer and a duplicate version of it built at Los Alamos under the supervision of Metropolis. Stanislav Ulam and Everett had already carried out computations on probability questions connected with the active assemblies of uranium and with neutron multiplications. The preferred name now and then is "branching processes," initially developed by David Hawkins. Edward Teller needed information about the progress of a thermonuclear reaction or burning in a mass of deuterium, or a deuterium and tritium mixture. There

were several proposals of ideas on how to initiate a thermonuclear reaction, using fission devices as ignition starters. An idea due to Gamow was called "the cat's tail," another was due to Edward Teller designated as " the womb idea," and another due to Stanislav Ulam designated as the "the spittoon idea." The work was involving the guessing of geometrical factors, intersections of solids, estimation of volumes, and estimation of the chances of points escaping. This gave doubts as to the possibility of success of the scheme suggested by Edward Teller for the Super. The geometry and density of the moving actual assembly had to be accounted for, as well as the fate of the neutrons and other particles moving through it and causing in turn more reactions in an autocatalytic way. These estimates had to be interspersed by stepwise calculations of the behavior of the actual hydrodynamic motions. The real times for the individual computational steps were short, each less than a "shake," where:

$$1 \text{ shake} = 10^{-8} \text{ [second]}.$$

The spatial subdivisions of the material assemblies very small in linear dimensions, and in special dimensions in the order of barns, where:

$$1 \text{ barn} = 10^{-24} \text{ [cm}^2\text{]}.$$

These early coupled Monte Carlo and Hydrodynamics calculations, necessitated the use of the MANIAC computer and its replica. The simulations results showed that the Edward Teller's configuration initially flared up, but then the whole assembly started to slowly cool down and fizzle.

Stanislav Ulam thought of a way to modify the whole approach by injecting a repetition of certain arrangements, which became known as the Ulam-Teller configuration. This breakthrough allowed the computer design of the first thermonuclear test and its successful testing in the Pacific as the Mike test. This idea or set of ideas is still shrouded with secrecy and classification considerations.

Owing to the secrecy surrounding the Super program, the exact character of that early work was not made known generally, and much confusion arose as to what Monte Carlo really was. When the general idea was explained, the statisticians suggested that this was nothing but an established statistical method called Model Sampling, which they have been doing for years. Stanislav Ulam and John von Neumann did not claim to have invented Model Sampling, but the essence of their contribution was:

1. The application of model sampling to neutron chain reactions,
2. The idea of doing the whole job on a digital computer,
3. The generation of the random numbers in the computer itself.

In fact, computers were invented and constructed to carry out such computations, so that modern day computers can be considered as a technological spin-off from the need to simulate chain reactions.

Harris and Herman Kahn systematically developed the method around 1948. At this time, Enrico Fermi, Metropolis and Stanislav Ulam estimated by Monte Carlo the eigenvalues of the Shrödinger equation.

Around 1970, the theory of computational complexity provided a more precise rationale for the use of the Monte Carlo method. It identified a class of problems for which the time to evaluate the exact solution to a problem within the class grows at least exponentially with the system dimension n. The question was whether the Monte Carlo method could estimate the solution to a problem in this intractable class to within a specified statistical accuracy in time bounded above by a polynomial in n-dimensional space. Many examples now support the ability of Monte Carlo to deal with these problems. For instance, Karp in 1985 shows this property by estimating the reliability in a planar multiterminal network with randomly failing edges, and Dryer in 1989 established it for the estimation of the volume of a convex body in an n-dimensional Euclidean space. In 1986 Broder and Jerrum and in 1988 Sinclair established this property in estimating the permanent of a matrix which is equivalent to the number of perfect matchings in a bipartite graph.

1.4. MONTE CARLO PROBLEM FORMULATION

The use of the Monte Carlo method is one of the best examples of the creative use of computers as a research tool. It draws on an extremely broad range of mathematical disciplines from probability theory to number theory, and from mathematical analysis to numerical analysis.

The Monte Carlo method depends on the use of probability theory to solve several categories of problems:

1. Problems that depend in their formulation in some way on probabilistic and random behavior.
2. Situations where physical experiments are impractical or too expensive.
3. Problems where an exact formula is difficult or impossible to derive using other known techniques.
4. Processes that consist of a long sequence of steps, each involving a probabilistic relation.

However, it is fundamentally different from probability theory. In probability theory, equations are derived from theoretical considerations based on randomness. In Monte Carlo theory, probabilistic concepts are used to find an answer to a physical problem that often has no relation to probability or randomness.

In a typical situation of a Monte Carlo problem formulation, designated as Direct Simulation Monte Carlo (DSMC):

1. The investigator has at hand a physical situation that he wants to study,
2. He cannot build an experiment to study the problem at hand,
3. He does not think he will even try to derive an equation describing the whole problem,
4. Even if he derives an equation, which is sometimes doubtful, not much information could be obtained from it,
5. He tries to identify a random process, which will give the same answers to his original situation, without having to derive an analytical solution,

6. He simulates the random process on a computer, and estimates on it the parameters of his original solution.

Consider the simple problem of neutron diffusion where a beam of neutrons impinges on a water tank. We need to calculate the fraction of neutrons getting through the tank without being absorbed or losing most of their energy. No formula could describe precisely the fate of all the neutrons. A Monte Carlo procedure would trace the life histories of a large sample of neutrons in the beam and their interaction with the water molecules and their hydrogen and oxygen constituents. The procedure would sample the source of neutrons and establish the first collision points in the tank. At each collision point, a decision is reached whether the particle has scattered, was absorbed or leaked from the system. If the particle was absorbed or leaked, a new particle is sampled from the source and its history is followed. If the particle has scattered, its energy loss is estimated and a new direction for its scattering is determined, then the position of its next interaction is determined. As the process is repeated, it creates a statistical population of neutrons in the tank that allows us to estimate its parameters, including the fraction leaking.

In an alternate situation that we can identify as Mathematical Monte Carlo Simulation (MMCS):

1. The physical situation under study has resulted in a mathematical equation, which due perhaps to multidimensionality, nonlinearity and complexity is very difficult or impossible to solve,
2. It cannot be solved by other methods without introducing drastic simplifications,
3. It cannot be solved in any reasonable amount of time using known techniques,
4. A probabilistic method, which would solve the mathematical equation and would yield a similar answer to the initial problem is sought, and simulated on a computer.

1.5 ERROR BEHAVIOR

For numerical methods aiming at estimating M points in an n-dimensional space, the Monte Carlo method is characterized by an absolute error that is inversely proportional to the square root of the number of sampled points N ≈ M, or:

$$|\varepsilon|_{MonteCarlo} \alpha_{\ell} \frac{1}{N^{\frac{1}{2}}}, \qquad (1)$$

whereas, absent any useful special structures, the absolute error in other numerical methods is inversely proportional to the n-th root of the number of points $\mathcal{N}$ as:

$$|\varepsilon|_{numerical} \alpha_{\ell} \frac{1}{M^{\frac{1}{n}}}. \qquad (2)$$

The implication is that for n>2, if we consider multidimensional systems, the Monte Carlo method becomes more favorable than other numerical methods.

1.6. PARAMETER ESTIMATION

In Monte Carlo simulations, statistical estimates are sought for the quantities of interest. The estimates are obtained by the repetitive playing of a probabilistic game. The game played is in the simplest case, but not necessarily so, an analog of the physical process of interest. The game is specified by a set of deterministic rules related to, and sets of probabilities governing the occurrences of the physical phenomena of interest.

Two distinctive features characterize the Monte Carlo method:

1. The simple structure of the computational algorithm, which simply involves the repetition of a numerical experiment for N times, then calculating mean values μ or mathematical expectations over the experiments,
2. A major advantage of the Monte Carlo methods is that the mean values of the quantities are associated with statistical error estimates. The absolute error can be calculated and it satisfies the relation:

$$\varepsilon = \frac{\sigma}{\sqrt{N}} \tag{3}$$

where σ is the square root of the variance or standard deviation of the sample process, and N is the number of conducted experiments.

Equation 3 suggests that reduced error bounds and a higher confidence in the obtained result can be obtained by increasing the number of experiments N in the denominator, or by decreasing the variance σ^2 of the statistical sampling process.

1.7. MONTE CARLO DECISION SUPPORT AND FORECASTING

Monte Carlo offers a future solution of what Sam Savage from Princeton University calls the "Flaw of Averages" problem in Decision Support systems and forecasting. Mean values that are used in today's decision support systems and forecasting can be very misleading. What engineers, scientists and economists have considered as exact until fairly recently are now recognized as mere mean values. In the real world and nature, the laws of probability and possibility theories prevail.

Even when given good data, people make bad decisions. They misunderstand, misinterpret and mismanage important problems. The cause is that our thought processes have bugs in them. To fix those bugs, decision support software latest incarnation generates thousands of scenarios covering real world contingencies. Software products such as XL Sim and @risk allow people to test their assumptions through random sampling and Monte Carlo simulations, turning spreadsheets into wind tunnels for testing hypotheses and designing good decisions.

To test their intuitive assumptions about average returns and average losses, executives managing plant capacity, investors with stock portfolios, employees with retirement's funds, would be running spreadsheet Monte Carlo simulations.

1.8. MONTE CARLO PERVASIVE COMPUTING

The Flaw of Averages also applies to Pervasive Computing. This is a future where computers are no more residing on a desktop or at a computer center, but are imbedded in the environment including workplaces, transportation means, and even clothing. For instance, a computer could control the temperature of clothing depending on the ambient environment. A computer could adjust the settings of light, communications gear, and climate control in an office, car or plane, depending on the present user of the facility.

At the MIT Sloan School of Management, Dan Ariely investigated an "electronic wallet." It would advise a user on, for instance, how to spend their money. Such a device would run a Monte Carlo simulation displaying how, based on a person's payment behavior, he would save a sum of $200 over the next few months by using one credit card or the other, check, or cash.

According to Michael Schrage: "… tomorrow's technologies will load the dice in favor of people not repeating the sort of silly statistical mistakes that lead to Nobel Prize winning research. And that will merit a prize of its own."

EXERCISE

1. Generate a plot comparing the behavior of the error as a function of the number of experiments N in a Monte Carlo computation.
Compare the previous result to the behavior of the error of a numerical computation if N $\approx$ M, for

a) One dimensional (n=1) problems,
b) Two dimensional (n=2) problems,
c) Three dimensional (n=3) problems.

Discuss the result and identify the region at which the use of a Monte Carlo computation is more favorable than a numerical one.