# Monte Carlo Simulations

## Background and history (edited from wikipedia.org)

Monte Carlo methods are a class of computational algorithms that rely on repeated random sampling to compute their results. Monte Carlo methods are often used in simulating physical and mathematical systems. These methods are most suited to calculation by a computer and tend to be used when it is infeasible to compute an exact result with a deterministic algorithm.This method is also used to complement the theoretical derivations.

Monte Carlo methods are especially useful for simulating systems with many coupled degrees of freedom, such as fluids, disordered materials, strongly coupled solids, and cellular structures (see cellular Potts model). They are used to model phenomena with significant uncertainty in inputs, such as the calculation of risk in business. They are widely used in mathematics, for example to evaluate multidimensional definite integrals with complicated boundary conditions. When Monte Carlo simulations have been applied in space exploration and oil exploration, their predictions of failures, cost overruns and schedule overruns are routinely better than human intuition or alternative "soft" methods.

The Monte Carlo method was coined in the 1940s by John von Neumann, Stanislaw Ulam and Nicholas Metropolis, while they were working on nuclear weapon projects in the Los Alamos National Laboratory. It was named in homage to Monte Carlo casino, a famous casino, where Ulam's uncle would often gamble away his money.

Before the Monte Carlo method was developed, simulations tested a previously understood deterministic problem and statistical sampling was used to estimate uncertainties in the simulations. Monte Carlo simulations invert this approach, solving deterministic problems using a probabilistic analog (see Simulated annealing).

In 1946, physicists at Los Alamos Scientific Laboratory were investigating radiation shielding and the distance that neutrons would likely travel through various materials. Despite having most of the necessary data, such as the average distance a neutron would travel in a substance before it collided with an atomic nucleus or how much energy the neutron was likely to give off following a collision, the problem could not be solved with analytical calculations.

Being secret, the work of von Neumann and Ulam required a code name. Von Neumann chose the name "Monte Carlo". The name is a reference to the Monte Carlo Casino in Monaco where Ulam's uncle would borrow money to gamble. Using lists of "truly" random numbers was extremely slow, von Neumann developed a form of making pseudorandom numbers, using the middle-square method. Though this method has been criticized as crude, von Neumann was aware of this: he justified it as being faster than any other method at his disposal, and also noted that when it went awry it did so obviously, unlike methods which could be subtly incorrect.

Monte Carlo methods were central to the simulations required for the Manhattan Project, though severely limited by the computational tools at the time. In the 1950s they were used at Los Alamos for early work relating to the development of the hydrogen bomb, and became popularized in the fields of physics, physical chemistry, and operations research. The Rand Corporation and the U.S. Air Force were two of the major organizations responsible for funding and disseminating information on Monte Carlo methods during this time, and they began to find a wide application in many different fields.

Uses of Monte Carlo methods require large amounts of random numbers, and it was their use that spurred the development of pseudorandom number generators, which were far quicker to use than the tables of random numbers that had been previously used for statistical sampling.

Monte Carlo methods vary, but tend to follow a particular pattern:

## Monte Carlo Simulations

1. Define a domain of possible inputs.
2. Generate inputs randomly from a probability distribution over the domain.
3. Perform a deterministic computation on the inputs.
4. Aggregate the results.

### Monte Carlo and random numbers

Monte Carlo simulation methods do not always require truly random numbers to be useful — while for some applications, such as primality testing, unpredictability is vital. Many of the most useful techniques use deterministic, pseudorandom sequences, making it easy to test and re-run simulations. The only quality usually necessary to make good simulations is for the pseudo-random sequence to appear "random enough" in a certain sense.

What this means depends on the application, but typically they should pass a series of statistical tests. Testing that the numbers are uniformly distributed or follow another desired distribution when a large enough number of elements of the sequence are considered is one of the simplest, and most common ones.

Sawilowsky lists the characteristics of a high quality Monte Carlo simulation:

- the (pseudo-random) number generator has certain characteristics (e. g., a long "period" before the sequence repeats)
- the (pseudo-random) number generator produces values that pass tests for randomness
- there are enough samples to ensure accurate results
- the proper sampling technique is used
- the algorithm used is valid for what is being modeled
- it simulates the phenomenon in question.

Pseudo-random number sampling algorithms are used to transform uniformly distributed pseudo-random numbers into numbers that are distributed according to a given probability distribution.

### Monte Carlo simulation versus "what if" scenarios

There are ways of using probabilities that are definitely not Monte Carlo simulations—for example, deterministic modeling using single-point estimates. Each uncertain variable within a model is assigned a "best guess" estimate. Scenarios (such as best, worst, or most likely case) for each input variable are chosen and the results recorded.

By contrast, Monte Carlo simulations sample probability distribution for each variable to produce hundreds or thousands of possible outcomes. The results are analyzed to get probabilities of different outcomes occurring. For example, a comparison of a spreadsheet cost construction model run using traditional "what if" scenarios, and then run again with Monte Carlo simulation and Triangular probability distributions shows that the Monte Carlo analysis has a narrower range than the "what if" analysis. This is because the "what if" analysis gives equal weight to all scenarios (see quantifying uncertainty in corporate finance).

### Applications

Monte Carlo methods are especially useful for simulating phenomena with significant uncertainty in inputs and systems with a large number of coupled degrees of freedom. Areas of application include:

- **Physical sciences -** Monte Carlo methods are very important in computational physics, physical chemistry, and related applied fields, and have diverse applications from complicated quantum chromodynamics calculations to designing heat shields and aerodynamic forms. In statistical physics Monte Carlo molecular modeling is an alternative to computational molecular dynamics, and Monte Carlo methods are used to compute statistical field theories of simple particle and polymer systems. Quantum Monte Carlo methods solve the many-body problem for quantum systems. In experimental particle physics, Monte Carlo methods are used for designing detectors, understanding their behavior and comparing experimental data to theory. In astrophysics, they are used to model the evolution of galaxies. Monte Carlo methods are also used in the ensemble models that form the basis of modern weather forecasting.

- **Engineering** - Monte Carlo methods are widely used in engineering for sensitivity analysis and quantitative probabilistic analysis in process design. The need arises from the interactive, co-linear and non-linear behavior of typical process simulations. For example,
    - in microelectronics engineering, Monte Carlo methods are applied to analyze correlated and uncorrelated variations in analog and digital integrated circuits. This enables designers to estimate realistic 3–sigma corners and effectively optimize circuit yields.
    - in geostatistics and geometallurgy, Monte Carlo methods underpin the design of mineral processing flowsheets and contribute to quantitative risk analysis.
    - impacts of pollution are simulated and diesel compared with petrol.
    - In autonomous robotics, Monte Carlo localization can be used to determine the position of a robot, it is often applied to stochastic filters such as the Kalman filter or Particle filter which form the heart of the SLAM ( simultaneous Localisation and Mapping ) algorithm.
- **Computational Biology -** Monte Carlo methods are used in computational biology, such for as Bayesian inference in phylogeny. Biological systems such as proteins membranes, images of cancer are being studied by means of computer simulations. The systems can be studied in the coarse-grained or ab initio frameworks depending on the desired accuracy. Computer simulations allow us to monitor the local environment of a particular molecule to see if some chemical reaction is happening for instance. We can also conduct thought experiments when the physical experiments are not feasible, for instance breaking bonds, introducing impurities at specific sites, changing the local/global structure, or introducing external fields.
- **Applied statistics** - In applied statistics, Monte Carlo methods are generally used for two purposes:
    - To compare competing statistics for small samples under realistic data conditions. Although Type I error and power properties of statistics can be calculated for data drawn from classical theoretical distributions (e.g., normal curve, Cauchy distribution) for asymptotic conditions (i. e, infinite sample size and infinitesimally small treatment effect), real data often do not have such distributions.
    - To provide implementations of hypothesis tests that are more efficient than exact tests such as permutation tests (which are often impossible to compute) while being more accurate than critical values for asymptotic distributions.
    - Monte Carlo methods are also a compromise between approximate randomization and permutation tests. An approximate randomization test is based on a specified subset of all permutations (which entails potentially enormous housekeeping of which permutations have been considered). The Monte Carlo approach is based on a specified number of randomly drawn permutations (exchanging a minor loss in precision if a permutation is drawn twice – or more frequently – for the efficiency of not having to track which permutations have already been selected).
- **Optimization** - Another powerful and very popular application for random numbers in numerical simulation is in numerical optimization. The problem is to minimize (or maximize) functions of some vector that often has a large number of dimensions. Many problems can be phrased in this way: for example, a computer chess program could be seen as trying to find the set of, say, 10 moves that produces the best evaluation function at the end. In the traveling salesman problem the goal is to minimize distance traveled. There are also applications to engineering design, such as multidisciplinary design optimization. Most Monte Carlo optimization methods are based on random walks. Essentially, the program moves randomly on a multi-dimensional surface, preferring moves that reduce the function, but sometimes moving "uphill".

- **Inverse problems** - Probabilistic formulation of inverse problems leads to the definition of a probability distribution in the model space. This probability distribution combines a priori information with new information obtained by measuring some observable parameters (data). As, in the general case, the theory linking data with model parameters is nonlinear, the a posteriori probability in the model space may not be easy to describe (it may be multimodal, some moments may not be defined, etc.). When analyzing an inverse problem, obtaining a maximum likelihood model is usually not sufficient, as we normally also wish to have information on the resolution power of the data. In the general case we may have a large number of model parameters, and an inspection of the marginal probability densities of interest may be impractical, or even useless. But it is possible to pseudorandomly generate a large collection of models according to the posterior probability distribution and to analyze and display the models in such a way that information on the relative likelihoods of model properties is conveyed to the spectator. This can be accomplished by means of an efficient Monte Carlo method, even in cases where no explicit formula for the a priori distribution is available. The best-known importance sampling method, the Metropolis algorithm, can be generalized, and this gives a method that allows analysis of (possibly highly nonlinear) inverse problems with complex a priori information and data with an arbitrary noise distribution.
- **Computational mathematics** - Monte Carlo methods are useful in many areas of computational mathematics, where a "lucky choice" can find the correct result. A classic example is Rabin's algorithm for primality testing: for any n which is not prime, a random x has at least a 75% chance of proving that n is not prime. Hence, if n is not prime, but x says that it might be, we have observed at most a 1-in-4 event. If 10 different random x say that "n is probably prime" when it is not, we have observed a one-in-a-million event. In general a Monte Carlo algorithm of this kind produces one correct answer with a guarantee n is composite, and x proves it so, but another one without, but with a guarantee of not getting this answer when it is wrong too often—in this case at most 25% of the time. See also Las Vegas algorithm for a related, but different, idea.
- **Design and visuals** - Monte Carlo methods have also proven efficient in solving coupled integral differential equations of radiation fields and energy transport, and thus these methods have been used in global illumination computations which produce photo-realistic images of virtual 3D models, with applications in video games, architecture, design, computer generated films, and cinematic special effects.
- **Finance and business** - Monte Carlo methods in finance are often used to calculate the value of companies, to evaluate investments in projects at a business unit or corporate level, or to evaluate financial derivatives. They can be used to model project schedules, where simulations aggregate estimates for worst-case, best-case, and most likely durations for each task to determine outcomes for the overall project.
- **Telecommunications -** When planning a wireless network, design must be proved to work for a wide variety of scenarios that depend mainly on the number of users, their locations and the services they want to use. Monte Carlo methods are typically used to generate these users and their states. The network performance is then evaluated and, if results are not satisfactory, the network design goes through an optimization process.