Explained: Monte Carlo simulations

May 17 2010, by Peter Dizikes

Monte Carlo

Speak to enough scientists, and you hear the words 'Monte Carlo' a lot. "We ran the Monte Carlos," a researcher will say. What does that mean?

The scientists are referring to Monte Carlo simulations, a statistical technique used to model probabilistic (or "stochastic") systems and establish the odds for a variety of outcomes. The concept was first popularized right after World War II, to study nuclear fission; mathematician Stanislaw Ulam coined the term in reference to an uncle who loved playing the odds at the Monte Carlo casino (then a world symbol of gambling, like Las Vegas today). Today there are multiple types of Monte Carlo simulations, used in fields from particle physics to engineering, finance and more.

To get a handle on a Monte Carlo simulation, first consider a scenario where we do not need one: to predict events in a simple, linear system. If you know the precise direction and velocity at which a shot put leaves an Olympic athlete's hand, you can use a linear equation to accurately forecast how far it will fly. This case is a deterministic one, in which identical initial conditions will always lead to the same outcome.

The world, however, is full of more complicated systems than a shot-put toss. In these cases, the complex interaction of many variables — or the inherently probabilistic nature of certain phenomena — rules out a definitive prediction. So a Monte Carlo simulation uses essentially random inputs (within realistic limits) to model the system and produce probable outcomes.

In the 1990s, for instance, the Environmental Protection Agency started using Monte Carlo simulations in its risk assessments. Suppose you want to analyze the overall health risks of smog in a city, but you know that smog levels vary among neighborhoods, and that people spend varying amounts of time outdoors. Given a range of values for each variable, a Monte Carlo simulation will randomly select a number within each range, and see how they combine — and repeat the process tens of thousands or even millions of times. No two iterations of the simulation might be identical, but collectively they build up a realistic picture of the population's smog exposure.

"In a deterministic simulation, you should get the same result every time you run it," explains MIT computer science professor John Guttag in his OpenCourseWare lecture on Monte Carlo simulations. However, Guttag adds, in "stochastic simulations, the answer will differ from run to run, because there's an element of randomness in it."

The aggregation of data makes it possible to identify, say, a median level of smog exposure. To be sure, Monte Carlo simulations are as good as their inputs; accurate empirical data would be necessary to produce realistic simulation results.

Provided by Massachusetts Institute of Technology