Monte Carlo Methods

Part I: Introduction

Spring Semester 2013/14

Department of Applied Mathematics, University of Crete

Instructor: Harmandaris Vagelis,

email: vagelis@tem.uoc.gr

Course web page: http://www.tem.uoc.gr/~vagelis/Courses/EM385/Monte_Carlo.html

What is Monte Carlo?

- The famous casino place

- But also ....

Monte Carlo Methods

- ❑ Monte Carlo numerical methods - Definition :

❖ Any computational method which solves a problem by generating suitable random numbers and observing that fraction of the numbers obeying some property or properties. The method is useful for obtaining numerical solutions to problems which are too complicated to solve analytically.

✓ It was named by S. Ulam, who in 1946 became the first mathematician to dignify this approach with a name, in honor of a relative having a propensity to gamble (Hoffman 1998, p. 239). Nicolas Metropolis also made important contributions to the development of such methods.

- ❑ They are also called "Monte Carlo Simulation" or "Stochastic Simulation" methods.

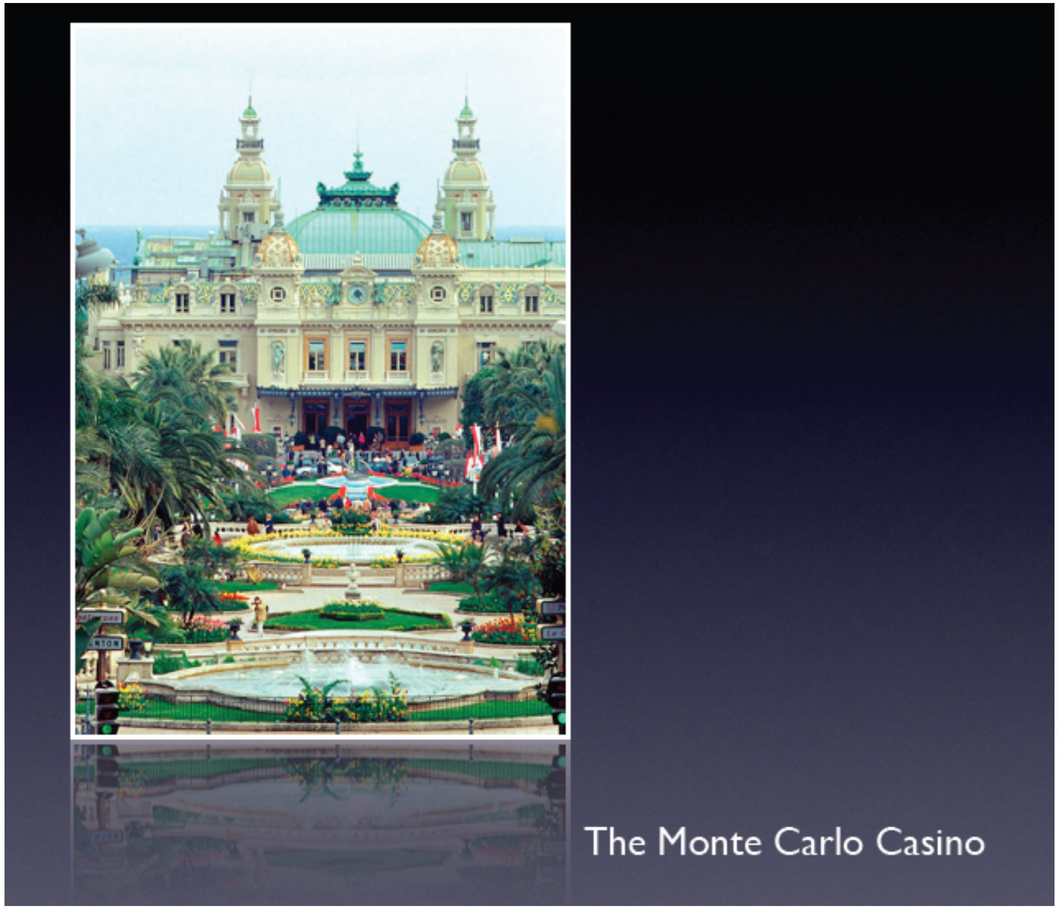
Monte Carlo Methods

❑ **ENIAC Los Alamos (1945-1955), from Stanislaw Ulam:**

➢ The first thoughts and attempts I made to practice [the Monte Carlo Method] were suggested by a question which occurred to me in 1946 as I was convalescing from an illness and playing solitaires. The question was what are the chances that a Canfield solitaire laid out with 52 cards will come out successfully? After spending a lot of time trying to estimate them by pure combinatorial calculations, I wondered whether a more practical method than "abstract thinking" might not be to lay it out say one hundred times and simply observe and count the number of successful plays. This was already possible to envisage with the beginning of the new era of fast computers, and I immediately thought of problems of neutron diffusion and other questions of mathematical physics, and more generally how to change processes described by certain differential equations into an equivalent form interpretable as a succession of random operations. Later [in 1946], I described the idea to John von Neumann, and we began to plan actual calculations.

Monte Carlo Methods

➢ 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 Mote Carlo Casino in Monaco where Ulam's uncle would borrow money to gamble!



Course Overview

- **Chapter 1 -** Introduction in Monte Carlo (MC): What is Monte Carlo? Deterministic vs Stochastic systems. Goals of the course.

- **Chapter 2 -** Basic Monte Carlo techniques: Random number generators. Numerical integration with MC. Optimization MC algorithms.

- **Chapter 3 -** Importance sampling: Monte Carlo importance sampling algorithms. Acceptance – rejection methods.

- **Chapter 4 -** Markovian chains: Metropolis – Hastings algorithms. Gibbs sampling.

- **Chapter 5 -** Special subjects: Convergence of MC algorithms. Comparing algorithms.

- **Chapter 6 -** Special subjects (graduate): Theoretical study of convergence of MC Markovian based algorithms. Multi-level techniques. Parallel MC. Models of natural systems.

Needed / Related Courses

- Needed courses - knowledge :

  ➢ Basic knowledge of probability theory, statistics. Introduction in probability (TEM-151).

  ➢ Programming language (C/C++ or Fortran 90 or Matlab)

  ➢ Stochastic methods I (TEM-261).

  ➢ Numerical methods for ordinary differential equations (TEM-291).

- Related courses:

  ➢ Stochastic methods II (TEM-262).

  ➢ Numerical methods for partial differential equations (TEM-292).

Overall Assessment

- ❑ Final exam (standard 1-2 hour examination): 70%, Lab/Coursework 30%.

or

- ❑ Final exam (standard 1-2 hour examination): 30%, Lab/Coursework 30%, Project: 40/%

✓ **Note:** The grade in the final examination and in the lab should be >= 5 (50%).

- ❑ The grades of the lab and the project are valid till September.

Teaching Course

➢ Lectures

➢ Lab / Coursework: Examples. Computer experiments using a programming language (e.g. C/C++, Fortran90, Matlab).

➢ Slides.

➢ 4-6 hours per week (2-4 theory + 2 coursework/lab).

Lab

Lab of the course involves:

➢ **Using a programming language or a software package (e.g. Matlab) for Monte Carlo techniques.**

➢ **Applying Monte Carlo algorithms in a series of different examples from statistics, natural sciences, economics, etc.**

❑ **During the semester 3-4 sets of coursework (problem sheets) will be given.**

➢ **The coursework is individual or in a small team of 2 persons.**

Project

- ❑ During the last 4-5 weeks of the semester a research project could be given.

- ❑ The project will concern the study and presentation of a research subject related with the theory and/or the application of Monte Carlo algorithms.

- ❑ The final project will be team based (small teams of 3-5 persons).

Goals of the Course

Upon completion of this course you should have:

➢ Acquired basic knowledge of the theory of Monte Carlo techniques.

➢ Understood Monte Carlo algorithms from the point of view of the numerical methods.

And be able to:

- ➢ Apply Monte Carlo methodologies in a variety of different problems.
- ➢ Use/modify Monte Carlo algorithms.
- ➢ Use stochastic simulation techniques.

Applications of Monte Carlo Algorithms

- ❖ Natural sciences: Statistical physics, chemistry, materials science, …
- ❖ Engineering,
- ❖ Economics,
- ❖ Mathematics: Applied Statistics,
- ❖ Astrophysics,
- ❖ Environmental models,
- ❖ Computational Biology, Biostatistics,
- ❖ Neural network models: modeling of brain,
- ❖ Many more …

Stochastic vs Deterministic

✓ Monte Carlo is a stochastic simulation technique:

❑ Deterministic system: A system in which the later states of the system follow from, or are determined (exactly) by, the earlier ones.

➢ No randomness.

➢ A deterministic model always produce the same output from a given starting condition or initial state.

❑ Stochastic system: A system in which the later states of the system are <u>not</u> determined (exactly) by the earlier ones.

➢ Involve randomness.

➢ There are many possible future states, for a given starting condition.

Stochastic vs Deterministic: Examples

- **Deterministic systems:** Classical mechanics, quantum mechanics, Navier - Stokes equations, … etc.

- **Stochastic systems**: Random walks, Brownian motion, stock market prices, biology (gene expression), … etc.

✓ **Note**: If a system is deterministic, this doesn't necessarily imply that later states of the system are predictable from a knowledge of the earlier ones. In this way, chaos is similar to a random system. For example, chaos has been termed "deterministic chaos" since, although it is governed by deterministic rules, its property of sensitive dependence on initial conditions makes a chaotic system, in practice, largely unpredictable.

History of Monte Carlo Methods

- (1733) Buffon's needle problem.
- (1812) Laplace suggests using Buffon's needle experiment to estimate π.
- (1946) ENIAC (Electronic Numerical Integrator And Computer) built.
- (1947) John von Neuman and Stanislaw Ulam propose a computer simulation to solve the problem of neutron diffusion in fissionable material.
- (1949) Metropolis and Ulam paper (Journal of the American Statistical Association).
- (1984) Gibbs sampler technique (Geman & Geman).

- From then onwards: continuously growing interest of statisticians in Monte Carlo methods.

A little bit of History: Buffon's Needle

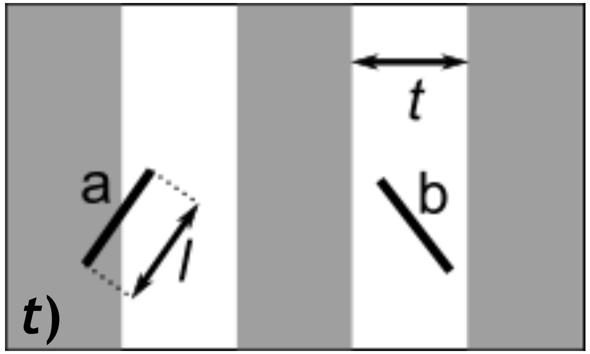
❖ First (?) application of Monte Carlo methods: Buffon's needle (Buffon 1777)

Georges-Louis Leclerc, Comte de Buffon
(7 September 1707 – 16 April 1788)

" Suppose we have a floor made of parallel strips of wood, each the same width, and we drop a needle onto the floor. What is the probability that the needle will lie across a line between two strips?"

➢ Solution (short needle $l < t$

$$P = \frac{2l}{\pi t}$$



➢ Home exercise:

A) Derive the above solution (short needle $l < t$)

B) What is the solution for large needle ($l > t$)?

Buffon's Needle

❖ Monte Carlo algorithm for the calculation of π:

A. "Throw" randomly n needles.
B. Check how many of them cut a line (suppose m).
C. Compute probability P_n through: $P_n = m/n$

$$\boxed{\pi = \lim_{n \to \infty} \frac{2l}{P_n t}}$$

The Best of the 20th Century: Top 10 Algorithms

❖ Dongarra and Sullivan, Computing in Science and Engineering, 2000, SIAM News.

❑ "Great algorithms are the poetry of computation," says Francis Sullivan of the Institute for Defense Analyses' Center for Computing Sciences in Bowie, Maryland. He and Jack Dongarra of the University of Tennessee and Oak Ridge National Laboratory have put together a sampling that might have made Robert Frost beam with pride--had the poet been a computer jock.

❑ Their list of 10 algorithms having "the greatest influence on the development and practice of science and engineering in the 20th century" appears in the January/February issue of Computing in Science & Engineering. If you use a computer, some of these algorithms are no doubt crunching your data as you read this.

The Best of the 20th Century: Top 10 Algorithms

The drum roll, please:

1946: The Metropolis Algorithm for Monte Carlo. Through the use of random processes, this algorithm offers an efficient way to stumble toward answers to problems that are too complicated to solve exactly.

1947: Simplex Method for Linear Programming. An elegant solution to a common problem in planning and decision-making.

1950: Krylov Subspace Iteration Method. A technique for rapidly solving the linear equations that abound in scientific computation.

1951: The Decompositional Approach to Matrix Computations. A suite of techniques for numerical linear algebra.

1957: The Fortran Optimizing Compiler. Turns high-level code into efficient computer-readable code.

1959: QR Algorithm for Computing Eigenvalues. Another crucial matrix operation made swift and practical.

1962: Quicksort Algorithms for Sorting. For the efficient handling of large databases.

1965: Fast Fourier Transform. Perhaps the most ubiquitous algorithm in use today, it breaks down waveforms (like sound) into periodic components.

1977: Integer Relation Detection. A fast method for spotting simple equations satisfied by collections of seemingly unrelated numbers.

1987: Fast Multipole Method. A breakthrough in dealing with the complexity of n-body calculations, applied in problems ranging from celestial mechanics to protein folding.

Bibliography

- *Monte Carlo Strategies in Scientific Computing*, J. Liu, Springer, New York, 2001.
- *Monte Carlo Statistical Methods*, C. Robert, G. Casella, Springer, New York, 2004.
- *Monte Carlo Simulation in Statistical Physics: An Introduction*, K. Binder, D. Heerrmann, Springer, New York, 2010.
- *A Guide to Monte Carlo Simulations in Statistical Physics*, D. Landau, K. Binder, Cambridge, New York, 2009.
- *Stochastic Methods: A Handbook for the Natural and Social Sciences*, C. Gardiner, Springer, New York, 2009.