

Lecture 16

Monte Carlo

T. Takahashi

Department of Mathematics and Statistics
University of New Mexico



Goals

We will answer these questions:

- What are Monte Carlo simulations?
- How does Monte Carlo integration work?
- What are its advantages?
- What are its disadvantages?



Average Values



Continuous Function $y = f(x)$

Definition: Average Value

The **average value** of a function $f : [a, b] \rightarrow \mathbb{R}$ is defined to be

$$av(f(x)) = f_{avg} = \frac{\int_a^b f(x) dx}{b - a}$$

- **Conceptually**, f_{avg} it is the value you get *on average* when you plug in a random number from a to b into $f(x)$
- **Geometrically**, $y = f_{avg}$ is a line that marks the average height of $y = f(x)$ over $[a, b]$



Example

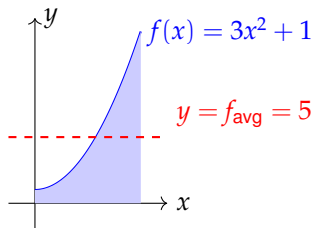
Example

Find the average value of $f(x) = 3x^2 + 1$ on $[0, 2]$.

- By definition

$$f_{avg} = \frac{\int_0^2 f(x) dx}{b-a} = \frac{(x^3 + x)|_0^2}{2-0} = \frac{10}{2} = 5$$

- Visual



Definition: Average Value for Scalar functions

Let $f(X)$ be a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ defined on $D \in \mathbb{R}^n$. Then the **average value** of f is

$$av(f(X)) := f_{avg} := \frac{\int_D f(X) dX}{\mu(D)},$$

where $\mu(D)$ is the **measure** of D .

- Conceptually, f_{avg} is the average value obtained when plugging in a random point on D
- The average value can be defined for *any* scalar-valued function
- The measure of a set is the length, area, or volume of the set
 - ▶ $\mu([3, 7]) = 7 - 3 = 4$
 - ▶ $\mu([a, b] \times [c, d]) = (b - a)(d - c)$
 - ▶ $\mu(\{x^2 + y^2 + z^2 < 1\}) = 4/3\pi r^3$



Practical Applications of Average Values

❶ **Snowfall:** $s(\text{location}, \text{day})$.

- ▶ $\text{location} = (x, y)$ coordinates over NM.
- ▶ $\text{day} = \text{day in a year}$.
- ▶ $av(s)$ = average snowfall over all of NM on any day of the year.

❷ **Option Pricing:** $f(S_1, S_2, \dots, S_n)$.

- ▶ S_i = price of asset i at maturity.
- ▶ n = number of underlying assets.
- ▶ $av(f)$ = expected payoff of the option, integrated over all asset price combinations.

❸ **Risk Assessment:** $f(X_1, X_2, \dots, X_n)$.

- ▶ X_i = uncertain parameter (e.g., material property, load condition).
- ▶ n = number of parameters.
- ▶ $av(f)$ = probability of system failure, integrated over all uncertain parameters.

❹ **Portfolio Optimization:** $f(w_1, w_2, \dots, w_n)$.

- ▶ w_i = weight of asset i in the portfolio.
- ▶ n = number of assets.
- ▶ $av(f)$ = expected portfolio return or risk, averaged over all possible allocations.



High-dimensional Integration

- For many practical applications we must approximate

$$\int_D f(X) dX$$

where X is a **high-dimensional variable**

- We have only discussed quadratures for when X is **one-dimensional**

$$\int_D f(X) dx = \int_a^b f(x) dx \approx \sum_i \alpha_i f(x_i)$$

Goal

We will learn a new easy way to integrate which works for *any* dimensions.



Random Variables



Discrete Random Variables

Definition: Discrete Random Variable

A **discrete random variable** X takes values in a *finite* set $\{x_0, x_1, \dots, x_n\}$ with probabilities $p_0, p_1, \dots, p_n$, such that $\sum_{i=0}^n p_i = 1$.

- X is the **random variable**
- Outcomes: $X_0, X_1, \dots, X_n$
- Probabilities $p_0, p_1, \dots, p_n$ with $\sum_{i=0}^n p_i = 1$
- We write $P(X = X_0) = p_0$

Example: Six-sided die

- X = a roll of a die
- Outcomes: $x_1 = 1, x_2 = 2, \dots, x_6 = 6$
- Probability $P(X = 4) = 1/6$



Expected Value and Variance

The Discrete Case

Definition: Discrete Statistics

The **expected value** of a discrete random variable x is

$$E[x] := \sum_{j=1}^n x_j p_j.$$

The **variance** of x is defined to be

$$\sigma^2[x] := E[(x - E[x])^2]$$

Example: Six-sided die

- $E[x] = (1 + 2 + \dots + 6)/6 = 3.5$
- $\sigma^2[x] = (1/6)(1 - 3.5)^2 + (1/6)(2 - 3.5)^2 + \dots + (1/6)(6 - 3.5)^2 = 2.916$



Continuous Random Variable

Definition: Continuous Random Variable

A **continuous random variable** X is a random quantity that can be an *infinite* number of values and the probability of taking on a value is represented by a distribution $\rho(x)$.

- Random variable range $X \in [a, b]$
- Probability: density function $\rho(x)$ with $\int_a^b \rho(x) dx = 1$
- Probability is expressed with inequalities e.g. $P(X_1 > X)$

Examples

- ▶ X = Height of a person
- ▶ X = Shelf life of a product
- ▶ X = Time it takes to solve a puzzle



Expected Value and Variance

The Continuous Case

Definition: Continuous Statistics

- The **expected value** of a continuous random variable X is

$$E[X] := \int_a^b x \rho(x) dx$$

- More generally, the expected value of $g(X)$ is

$$E[g(X)] := \int_a^b g(x) \rho(x) dx$$

- The **variance** of X is defined to be

$$\sigma^2[X] := E[(x - E[x])^2] = \int_a^b (x - E[x])^2 \rho(x) dx$$



Uniform Random Variable

Definition: Uniform Random Variable

A **uniform random variable** is a continuous variable that takes on a random number between a and b with equal probability.

- By definition, if X is uniform then the density function must be

$$\rho(x) = \frac{1}{b - a}$$

Example

Let $a = 1$, $b = 3$ (standard uniform distribution). Compute $P(1.2 \leq X \leq 1.5)$:

$$P(0.2 \leq X \leq 0.5) = \int_{1.2}^{1.5} \frac{1}{3 - 1} dx = 0.15 = 15\%$$



Monte Carlo Simulation



Rolling a Die

- We may estimate the expected value of X , by drawing samples and taking the average

$$E[x] = \frac{1}{N} \sum_{j=1}^N x_j$$

- Example - let X be the outcome of a six-sided die

Estimated Average

- ▶ 3 rolls: 3, 1, 6 $\rightarrow E[x] \approx (3 + 1 + 6)/3 = 3.33$
- ▶ 9 rolls: 3, 1, 6, 2, 5, 3, 4, 6, 2 $\rightarrow E[x] \approx (3 + 1 + 6 + 2 + 5 + 3 + 4 + 6 + 2)/9 = 3.51$
- As the number of rolls N increases, we approach the true value $E[X] = 3.5$



Definition: Monte Carlo Simulation

Monte Carlo simulation is a mathematical technique that uses repeated random sampling to estimate an outcome of a system with uncertain variables.

- Key Components
 - Use knowledge of distribution of random variables in the sampling
 - Increase repetition to reduce error and improve accuracy.
- Applications:
 - Estimating integrals (e.g., high-dimensional).
 - Risk analysis and financial modeling.
 - Simulating physical systems (e.g., particle dynamics, climate models).
- Examples:
 - Dice rolls
 - Coin flipping



Running a Simulation in Matlab

- To simulate a dice roll, do

$$X = \text{randi}(6)$$

- To simulate a uniform random variable from 0 to 1, do

$$X = \text{rand}$$

- To simulate a coin toss (heads=1, tails=0) do

$$X = (\text{rand} > 0.5)$$

- To simulate a uniform random variable from a to b , do

$$X = (b-a)*\text{rand} + a$$


Simple Monte Carlo Example

Example

On average, how many coin flips would it take to land 3 heads in a row?

```
1 flips_needed = zeros(1,100000);  
2 for trial = 1:100000  
3     consecutive_heads = 0;  
4     flips = 0;  
5     while consecutive_heads < 3  
6         flips = flips + 1;  
7         if rand() < 0.5 % Simulate a coin flip  
8             consecutive_heads = consecutive_heads + 1;  
9         else  
10            consecutive_heads = 0; % Reset if Tails  
11        end  
12    end  
13    flips_needed(trial) = flips;  
14 end  
15 average_flips = mean(flips_needed)
```

Output: average_flips = 14.0289

The Real Answer?

Example

On average, how many coin flips would it take to land 3 heads in a row?

We model progress toward 3 consecutive heads using states:

S_0 : no streak, S_1 : 1 head, S_2 : 2 heads.

Let E_0, E_1, E_2 be the expected additional flips to reach 3 heads in a row.

- From S_0 :

$$E_0 = 1 + \frac{1}{2}E_1 + \frac{1}{2}E_0 \Rightarrow E_0 = 2 + E_1.$$

- From S_1 :

$$E_1 = 1 + \frac{1}{2}E_2 + \frac{1}{2}E_0.$$

- From S_2 :

$$E_2 = 1 + \frac{1}{2}E_0.$$

Solving the system gives: $E_0 = 14$. Monte Carlo would eventually give the same answer.



Average Value Example

Example

On average, what is the value of $f(x) = 3x^2 + 1$ for a uniform random number $x \in [0, 2]$?

```
1 fvals = zeros(1,1000000);  
2 f=@(x) 3*x^2 + 1;  
3 for trial = 1:1000000  
4     X = 2*rand; % random number on [0,2]  
5     fvals(trial) = f(X);  
6 end  
7 average_f = mean(fvals)
```

Output: average_f = 5.0005

Notice how the output is close to $av(f) = 5$



Question

Why does Monte Carlo simulation work?

- As $N \rightarrow \infty$, the error between the estimate and the true value is statistically zero.
- That is, the estimate will converge to the correct value.
- The probability $P(\cdot)$ is

$$P \left(E[x] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N x_i \right) = 1$$

- This theoretical concept is called the **Law of Large Numbers**
- Increasing the number of trials eventually makes the error go to zero



Monte Carlo Integration



Estimating $E[g(X)]$

- Suppose X is a uniform random variable from a to b
- Let $g(x)$ be some continuous real-value function $g : \mathbb{R} \rightarrow \mathbb{R}$
- By definition, the average value of $g(X)$ is exactly

$$\begin{aligned} E[g(X)] &= \int_a^b g(x) \rho(x) dx \\ &= \int_a^b g(x) \left(\frac{1}{b-a} \right) dx \\ &= \frac{\int_a^b g(x) dx}{b-a} \end{aligned}$$

- We can approximate $E[g(X)]$ with a Monte Carlo simulation

$$E[g(X)] \approx \frac{1}{N} \sum_{i=1}^N g(x_i)$$



Connection Between Average Value and Integrals

- The average value of a function $E[f(X)]$ where X is uniform on $[a, b]$ is

$$E[f(X)] = \frac{1}{b-a} \int_a^b f(x) dx$$

- This is exactly equal to the average value of $f(x)$ for $a \leq x \leq b$
- Therefore, we have

$$\int_a^b f(x) dx = (b-a)E[f(X)] = (b-a)f_{avg}$$

- Using Monte Carlo simulation we have

$$\int_a^b f(x) dx \approx (b-a) \frac{1}{N} \sum_{i=1}^N f(x_i)$$



Monte Carlo Integration

- Using the Monte Carlo simulation we have the approximation

$$\frac{\int_a^b g(x) dx}{b-a} \approx \frac{1}{N} \sum_{i=1}^N g(x_i)$$

- That means we have the following formula for the integral of $g(x)$:

Monte Carlo Integral Approximation

$$\int_a^b g(x) dx \approx \frac{b-a}{N} \sum_{i=1}^N g(x_i),$$

where x_i are random uniform variables on $[a, b]$.



Monte Carlo Integration: Algorithm

- For uniform random numbers $x_1, \dots, x_n$, we have the numerical integration algorithm

$$\int_0^1 f(x) dx \approx \frac{1}{N} \sum_{i=1}^N f(x_i)$$

- The implementation in Matlab is very simple:

```
1 n=100;  
2 x=rand(n,1); % generate a vector of random numbers at once  
3 a=f(x);  
4 s=sum(a)/n;
```

Method: Monte-Carlo Integration

- Given:** Function $f(x)$, interval $[0, 1]$, N random points on $[0, 1]$
- Output:** Approximation to $\int_0^1 f(x) dx$
- Idea:** We can keep taking samples. The law of large numbers tells us this method converges.

Example for 1D

Example

Use Monte Carlo integration to approximate $f(x) = 3x^2 + 1$ for $x \in [0, 2]$. For the random variables X , use the values

0.1, 0.2, 1.2, 1.5.

- The approximation is

$$\int_0^2 f(x) dx \approx \frac{2-0}{4} \sum_{i=1}^4 f(x_i)$$

- Using the random numbers given, we have

$$\begin{aligned} \int_0^2 f(x) dx &\approx \frac{1}{2} \sum_{i=1}^4 (3(x_i)^2 + 1) \\ &= \frac{1}{2} [3(0.1)^2 + 3(0.2)^2 + 3(1.2)^2 + 3(1.5)^2 + 4] = 7.61 \end{aligned}$$



Two Dimensions



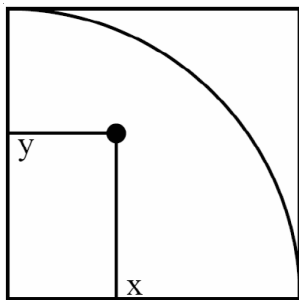
Two Dimensional Problem

- Consider integrating the function of two variables

$$f(x, y) = \begin{cases} 1 & (x, y) \in \text{circle} \\ 0 & \text{else} \end{cases}$$

- Let the domain of integration be the unit square $[0, 1]^2$
- Then the exact integral is just the area of the quarter circle

$$\int_0^1 \int_0^1 f(x, y) dx dy = A = \frac{\pi}{4}$$



Approximate π

Goal

Knowing that

$$\int_0^1 \int_0^1 f(x, y) dx dy = \frac{\pi}{4}$$

use Monte Carlo integration to approximate the value of π .

How?

- Generate multiple random variables X that are uniform in $[0, 1]^2$
- Plug values of X into f and sum them
- Divide sum by total number of points
- Multiply by 4 to get approximation of π

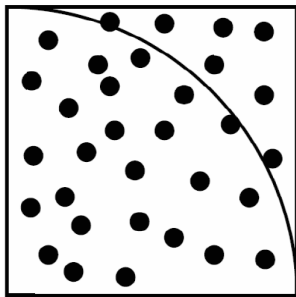
$$\pi \approx \frac{4}{N} \sum_{i=1}^N f(x_i, y_i)$$



The Dartboard Analogy

- 1 Throw N darts randomly in $[0, 1]^2$
- 2 Count the number of darts that land inside $x^2 + y^2 < 1$
- 3 Divide total by N , then multiply by 4

$$\pi \approx 4 \frac{\text{number of darts landed inside}}{\text{total darts thrown}}$$



Numerical Result

```
1 Input: N, seed
2 rng(seed) % set the seed
3 call rand in 2D to generate  $(x_i, y_i)$ 
4 for i=1:N
5      $sum = sum + f(x_i, y_i)$ 
6 end
7  $sum = 4 * sum / N$ 
```

```
1 Usage: monte_carlo_pi( niter, seed );
2
3 >>> monte_carlo_pi( 1000, 0 )
4 ... estimate of pi is 3.1320000000000000
5
6 >>> monte_carlo_pi( 10000000, 0 );
7 ... estimate of pi is 3.14243...
```

- $N = 1000$ gives us 2 digits
- $N = 10,000,000$ gives us 4 digits



Monte Carlo Convergence Rate

The expected value of the error is $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$

- We are guaranteed this error by the law of large numbers
- Convergence does not depend on dimension

General Convergence

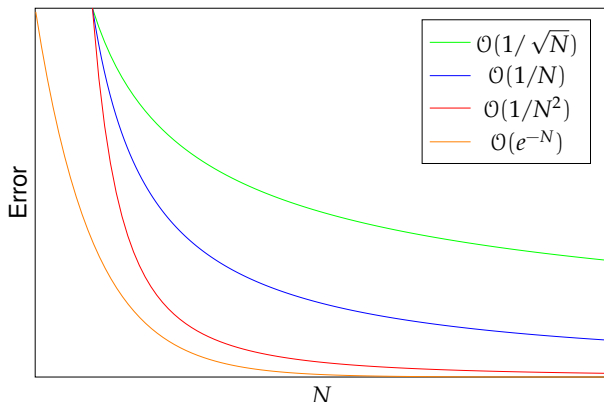
The expected value of the error in general for Monte Carlo integration is

$$\mathcal{O}\left(\frac{\sigma}{\sqrt{N}}\right)$$

where σ is the standard deviation (square-root of the variance).



Disadvantage of Monte Carlo



- The main disadvantage of Monte Carlo is its slow convergence rate



Multidimensional extensions

Why do we care about a method that converges as slowly as $\mathcal{O}(\frac{1}{\sqrt{N}})$?

- High-dimensional quadratures are difficult to derive
- Many of the more sophisticated integration methods scale poorly with dimension
- Let D be a difficult domain to integrate, e.g., complex geometry or high dimensionality
- Expected value

$$E[g(x)] = \int_{x \in D} g(x) \rho(x) dx \approx \frac{1}{N} \sum g(x_i)$$

Key Idea

- ▶ Monte-Carlo is **robust** and **easy to implement**
- ▶ Then convergence is **unaffected by the dimensionality** of the problem

Monte Carlo Summary

Method: Monte Carlo Integration

- **Given:** $f : \mathbb{R}^n \rightarrow \mathbb{R}$
- **Output:** Approximation to $\int_D f(X) dX$
- **Idea:** Use random sampling to find the average value of f

Disadvantages

- Slow convergence rate $\mathcal{O}(N^{-1/2})$

Advantages

- No increase in implementation difficulty w.r.t. dimension of X
- No slow down in convergence rate w.r.t. dimension of X

