DATA252/DATA551: Modeling and Simulation

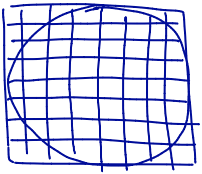
Lecture 8: Monte Carlo methods: simple examples

April 6, 2020

Introduction

- Monte Carlo methods refer to a very broad class of methods that rely on using random sampling to solve otherwise difficult problems.
- Example: bootstrapping; estimating π; Monte Carlo tree search (MCTS)[1]

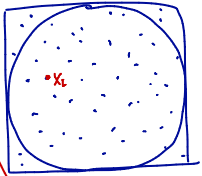
None - monte carlo method



estimate the area of the circle

$X_1 \cdots X_m \sim$ Unif on $\mathbb{R}^2$

Monte - Carlo Method



$$\frac{\#\text{dots in circle}}{\#\text{dots total}} = \frac{\text{area of circle}}{\text{area of square}}$$

$X_i = \begin{pmatrix} X_i^1 \\ X_i^2 \end{pmatrix} \to (x\text{-}y)$ coordinate

- Today, we'll use Monte Carlo (i.e., simulation) for two simple tasks: 1) integration and 2) calculating p-value

[1] See https://medium.com/@quasimik/monte-carlo-tree-search-applied-to-letterpress-34f41c86e238 for a nice introduction

Monte Carlo Integration

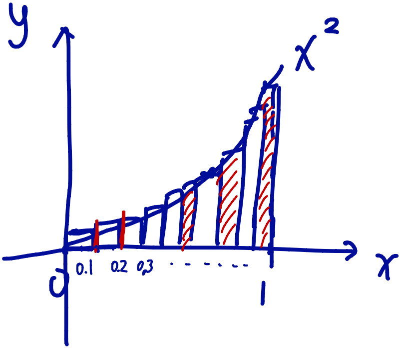
Monte Carlo Integration is one of the most commonly used Monte Carlo methods.

Ex1: Calculate the integral $\int_0^1 x^2 dx$.

Math: $\frac{1}{3}x^3\Big]_0^1 = \frac{1}{3}$

none-Monte Carlo approximation method:

$$\left[(0.1)^2 + (0.2)^2 + \cdots + (0.9)^2 + 1^2 \right] \cdot \frac{1}{10}$$



Monte Carlo approximation method:

instead of fixed grid $0, 0.1, 0.2, 0.3, \cdots, 1$

randomly generate numbers from $[0, 1]$: $X_1 \cdots X_m$

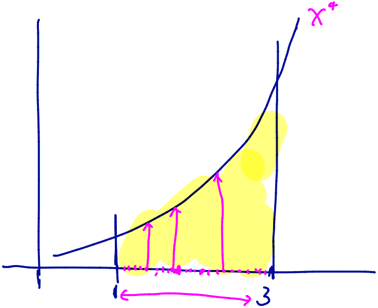
$$(X_1^2 + X_2^2 + \cdots X_m^2) \cdot \frac{1}{m}$$

Monte Carlo Integration

Ex2: Calculate the integral $\int_1^3 x^4 dx$.

(Math: $\left.\frac{x^5}{5}\right]_1^3 = \frac{3^5}{5} - \frac{1}{5} = 48.4$)

optional



① random sample points on [1, 3]

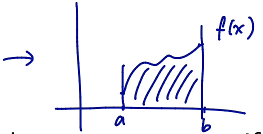
② calculate height at each point : $(\text{data})^4$

③ calculate height * 2 / # of points you sampled

$= \text{mean}(\text{data}^4) \times 2$

Monte Carlo Integration

Algorithm for estimating $\int_a^b f(x)dx$:



Step 1 Make a large number of random draws $x_1, \ldots, x_m$ uniformly from $[a, b]$.

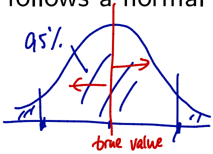
Step 2 Calculate $f(x_1), \ldots, f(x_m)$.

Step 3 Calculate $(b-a) \cdot \bar{f}$ (where $\bar{f}$ is the average of $f(x_1), \ldots, f(x_m)$) as an estimate of $\int_a^b f(x)dx$.

Exercise: use this algorithm with $\underline{m = 1000}$ to estimate $\boxed{\int_0^2 \underbrace{x^3 - x^2}_{f(x)} dx}$. Run it twice and report both estimates.

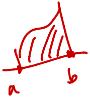
Discussions

- The Monte Carlo estimates are random, and the randomness follows a normal distribution due to Central Limit Theorem.



If you generate a lot of estimates, you can construct confidence intervals for the true value.

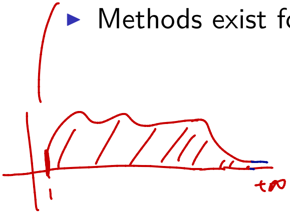
Or you can use normal distr. to get the CI.



- Extensive efforts have been made to reduce the standard error of the Monte Carlo estimates.

$\int_a^b$

- Methods exist for evaluating integrals that are not defined on bounded intervals $[a, b]$.
- Methods exist for evaluating higher dimensional integrals.



Hypothesis testing

A classic example in statistics, due to R. A. Fisher: *lady tasting tea*

- T: adding tea first; M: adding milk first
- A lady claims that she can tell the difference by tasting
- 8 cups are prepared: T T T T M M M M
- The lady is asked to pick out 4 that she believes was made by adding tea first (T)

Out of the 4 she picks, how many needs to be correct for you to believe that she actually has the ability to tell the difference?

Statistical terms you might have learned before...

- Null hypothesis H_0: The lady is just guessing (i.e., cannot tell the difference)
- Alternative hypothesis H_1: The lady can tell the difference
- Test statistic: The number she gets correct (0, 1, 2, 3, or 4)
- Reference distribution: Distribution of the test statistic under the null hypothesis

If she is just guessing...

0	1	2	3	4
1%	24%	50%	24%	1%

(based on simulation)

- p-value: Probability that the test statistic is as extreme, or more extreme, than the observed value, under the null hypothesis.

Let's say she get 3 out of 4 correct;

then the p-value = 24% + 1% = 25% or 0.25

(recall: cutoff for p-value is usually 0.05)

Large p-value ⇒ don't have enough evidence to reject the null (she's guessing)

⇒ don't have enough evidence to say that she can tell the difference.

Let's say she get 4 out of 4 correct:

If she's just guessing, there's only 1% chance she gets 4 out 4 correct, so she's probably not guessing.

(we have enough evidence to reject the null that she is guessing, p-value = 0.01)

Coin flipping example

10 flips

Flip a coin, and we get HHHTHHHHTH. Do you think we have enough evidence to say that this is a biased coin? Yes or No

- Null hypothesis: Fair coin ($p=0.5$)
- Alternative hypothesis: $p \neq 0.5$
- Test statistic: # of heads out of 10 flips
- Reference distribution: If $p=0.5$, what is the distribution of #head out of 10 flips? (binomial)

"as extreme or more extreme"

0	1	2	3	4	5	6	7	8	9	10
0.0006	0.0092	0.0440	0.1148	0.2150	0.2492	0.1980	0.1142	0.0452	0.0086	0.0012

observed

"more extreme"

- p-value:

In reality: we get 8 heads out of 10 flips.

$$\text{p-value} = 0.0006 + 0.0092 + 0.0440 + 0.0452 + 0.0086 + 0.0012 > 0.05$$

$$= 0.109 \approx 11\%$$

We do not have enough evidence to say this is a biased coin.

If we get 9 heads out of 10 flips, p-value = 0.0006 + 0.0092 + 0.0086 + 0.0012 = 0.02 or 2%

small p-value (< 0.05), we have enough evidence to say it's a biased coin.

Discussions

see wikipedia　　　binomial

- In both of “lady tasting tea" and coin flipping examples, the reference distribution is actually easy to derive without simulation. Simulation is more useful with more difficult test statistics and reference distributions.

- If you observe 9 heads out of 10 ~~heads~~ flips, you would reject the null of $p = 0.5$. Does that mean the coin is definitely unfair?

  No.

- If you observe 8 heads out of 10 ~~heads~~ flips, you would “accept" the null of $p = 0.5$. Does that mean the coin is definitely fair?

  No. "fail to reject" the null; usually we don't say: "we have proved the null" or "the null is correct"

  Hypothesis testing is never about "proving" a hypothesis.