Monte Carlo Simulations

By Naufil Sakran

Contents

- What is Monte Carlo Simulations
- The working of Monte Carlo Simulations
- Determining the probability of a coin toss using Monte Carlo Simulations
- Roulette game
- Solving the Buffon Needle Problem using Monte Carlo Simulations

History!

2+2=4

- The concept was invented by the Polish American Mathematician Stanislaw Ulam in 1946, while he was working on thermonuclear weapons.
- During his ill days, he mostly spend his time playing solitaire. Being a Mathematician he wanted to know the probability of him winning the game. He tried to find the probability using combinatorial calculations but he failed.
- An idea occurred to him that why don’t he simulate the game on a computer and compute its probability. He asked John Von Neumann to simulate the game on computer and thus Monte Carlo Simulation was born.
- Since the communication between them was secretive, they used the code name Monte Carlo for their project and hence the method was name Monte Carlo Simulation.
- Later the idea of it, was used in the development hydrogen bombs.

What is Simulations ?

- Simulation is the process of designing a model of a real system
- Then we conduct experiments with the model for the purpose of understanding the behaviour of the operation of the system.
- We duplicate the Original system and use it to understand the behaviour of the system

What is Monte Carlo Simulations Technique ?

A method of estimating the value of an unknown quantity using the principles of inferential statistics.

Technique!!!

- We establish the probability distribution of the system.
- We simulate it on a computer, and run it large number of times i.e. we draw **large** number of random samples.
- Random samples tends to exhibit the **same properties** of the population from which it is drawn.
- We then do some statistics on it and make inference about the population.

Law of large Numbers

- Suppose that $X_1, X_2, \ldots, X_n$ form a **random sample** from a distribution from which the mean is μ and for which the variance is finite. Let $\overline{X_n}$ denote the sample mean. Then

$$\lim_{n\to\infty} P(|\overline{X_n} - \mu| < \varepsilon) = 1$$

- As the sample size increases, the sample mean approaches the population mean with probability 1.

Secret behind the success of Monte Carlo Simulation

- The Law of Large Numbers plays a significant role in success of Monte Carlo Simulations.
- The larger the size of our random samples the less variability we get in the distribution of the sample means i.e. the **expectation of the sample means become closer to the true population mean** as a consequence of law of large numbers.
- Our inference about the population become more accurate as the number of random samples increases
- Regression to the mean

Regression to the mean

- Following an extreme **random** event, the next random event is likely to be less extreme
- Francis Galton,1885 proved statistically(with data) that if someone’s parent are both taller than average, its likely that the child would be shorter than the parents (closer to average). Following an extreme event, parents being unusually tall, the next random event is likely to be less extreme.
- The next random event is likely to move toward the average.

Determining the probability of a coin toss using Monte Carlo Simulations

- Suppose we flip a coin once and we get a head. What would be our guess of the next flip. (not sure)

- Suppose we flip a coin twice and we get 2 heads. Our guess of the next flip would be vey little inclined towards head but still we would be not very sure.

- Suppose we flip a coin hundred times.
- Suppose we get 100 heads and 0 tails (rare event)
- So our best assumption that the next flip will be head is 100%. We will be pretty confident in our answer.
- Suppose we get 70 heads and 30 tails
- We will be 70% confident in our guess that the next flip will be head.
- Suppose we get 40 heads and 60 tails
- Our confidence of head in the next flip will be 40%.

Why the difference in confidence ?

Confidence in our estimate depends upon two things:

- Variance of samples
- Size of samples

As the variance grows, we need larger samples to have the same degree of confidence
We now simulate it on computer and see the results.

European Roulette

- We choose a number from 1-36
- The wheel is then spun
- A ball is dropped
- If the ball lands on our chosen number, we win
- If we bet $1 on a number, if we win we get $35 dollars + our $1.
- If we lose, we lose our $1
- If the ball lands on 0, the casino wins

“The TABLE always win”

- For a fair game, the expectation of return must be 0. Which means we have equal probability of gaining money or losing money.
- We bet \$1.
- Fair Roulette
- $E(X) = 35 * \frac{1}{36} + (-1) * \frac{35}{36} = 0 \quad \therefore 0\ is\ not\ included$
- European Roulette
- $E(X) = 35 * \frac{1}{37} + (-1) * \frac{36}{37} = -\frac{1}{37} = -0.027 \quad \therefore 0\ is\ included$
- We use Monte Carlo Simulation to prove our claim.

Results

95% Confidence Interval of 100 trials of 100 simulations:

- **-0.1358 < Expected Return of Fair Roulette < 0.1374, Point Estimate=0.0824**
- **-0.1895 < Expected Return of European Roulette < 0.0183, Point Estimate= -0.0856**

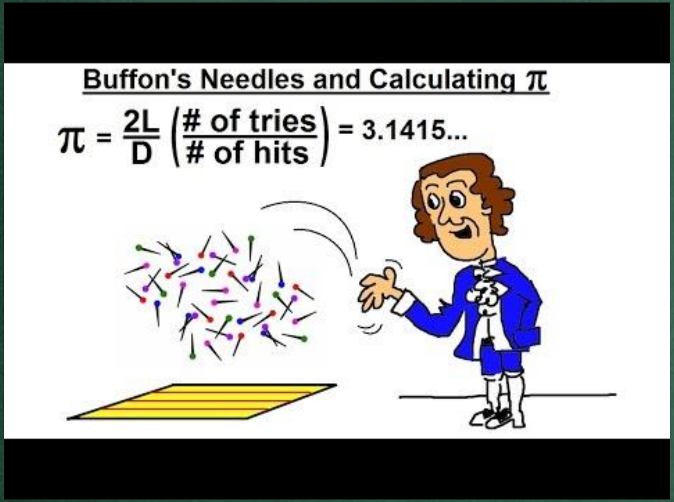
95% Confidence Interval of 1000 trials of 100 simulations:

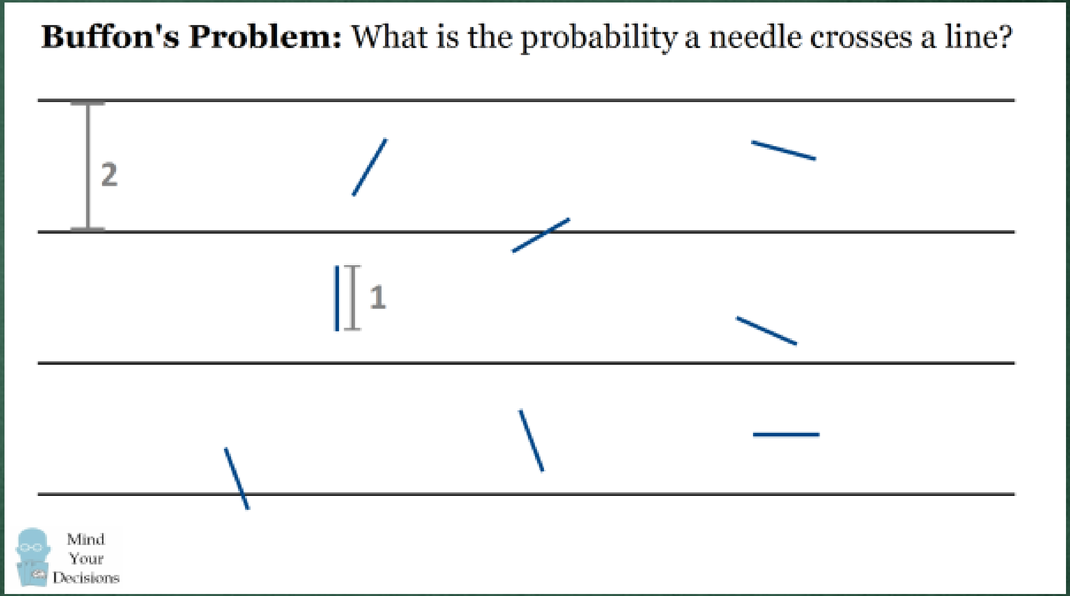
- 0.0009 < Expected Return of Fair Roulette < 0.0252**, Point Estimate=0.0130**
- -0.0220 < Expected Return of European Roulette < -0.0009**, Point Estimate= -0.0144**

95% Confidence Interval of 100 trials of 10000 simulations:

- **-0.0024 < Expected Return of Fair Roulette < 0.0012, Point Estimate=0.0023**
- **-0.0276 < Expected Return of European Roulette < -0.0255, Point Estimate= -0.0266**

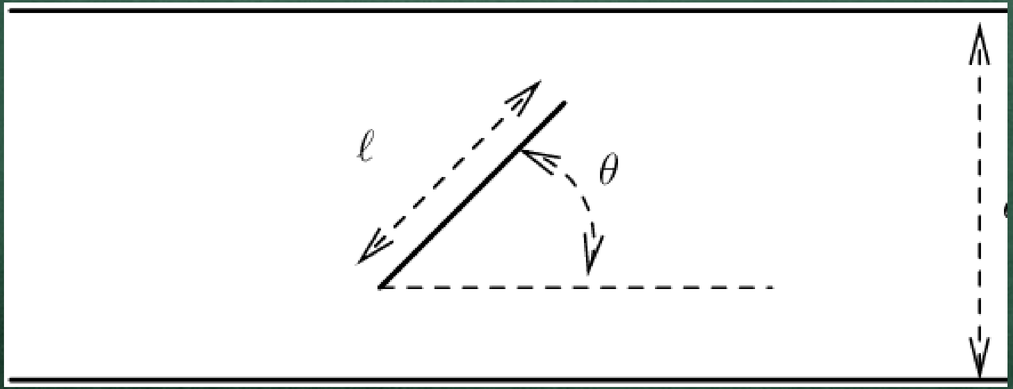
Buffon Needle Problem





Model

- Let x denote the distance from the midpoint of the needle to the nearest line
- Let θ denote the acute angle between the extended end of the needle and the line
- X: $f_x(x) = \frac{2}{d}, \quad where\ 0 \le x \le \frac{d}{2} \quad \therefore\ uniformly\ distributed, \frac{2}{d}\ is\ the\ normalizer$
- $\theta : f_\theta(\theta) = \frac{2}{\pi}, \quad where\ 0 \le \theta \le \pi/2 \ \therefore\ uniformly\ distributed, \frac{2}{\pi}\ is\ the\ normalizer$



In mathematical language, the needle will intersect the line when

$$X \le \frac{l}{2}\sin(\theta)$$

Joint probability distribution

$$f_{x,\theta}(x,\theta) = \frac{4}{\pi d} \quad where\ 0 \le x \le \frac{d}{2}\ ,0 \le \theta \le \frac{\pi}{2}$$

So,

$$P\left(X \le \frac{l}{2} sin\theta\right) = \int_0^{\frac{\pi}{2}} \int_0^{\frac{l}{2}sin\theta} \frac{4}{\pi d}\, dx d\theta = \frac{\mathbf{2l}}{\boldsymbol{\pi d}}$$

Estimating using Monte Carlo Simulation

- We consider the case where the distance between the lines is 2cm and the length of the needle is 1cm.

d=2cm and ℓ=1cm

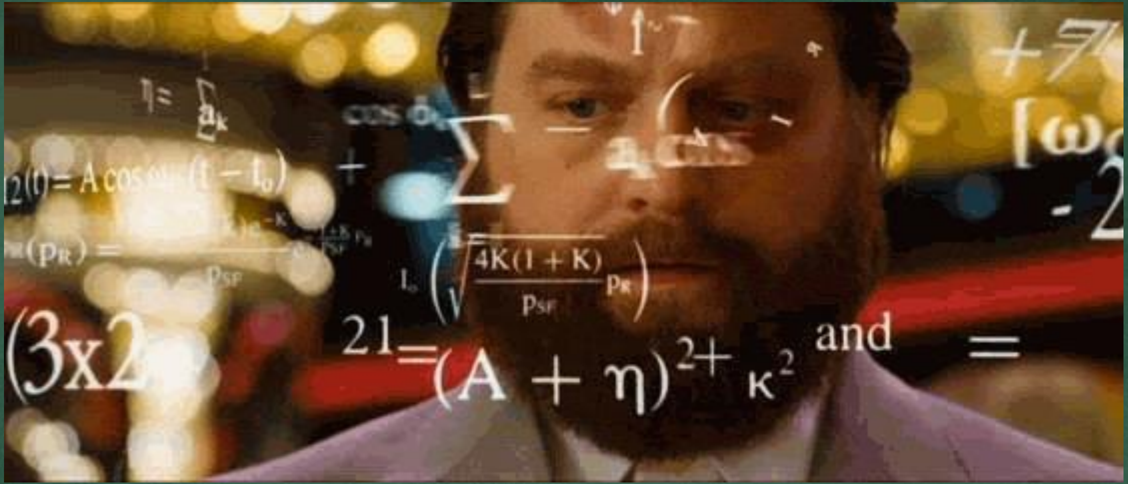
So we should have

$$P\left(X \le \frac{l}{2} sin\theta\right) \approx \frac{1}{\pi}$$

- Hence we could approximate the value of π. As our sample size tends to infinity, our estimate tends to the value $\frac{1}{\pi}$.

References

- Lecture 6: Monte Carlo Simulation by Prof. John Guttag, MIT, website: ocw.mit.edu
- A First Course in Probability by Sheldon Ross, p243-244 , p400, p438
- Probability and Statistics by Morris H. DeGroot and Mark J.Schervish, p352-753, p787-790



THANK YOU!

Any questions?

You can find me at naufilsakran@live.com