A Review of Applied Probability

Law of Large Numbers, Monte-Carlo Simulation

Dr. Bahman Honari

Autumn 2022

Introduction

- Instead of using mathematical and algebraic solution for probability questions, we now have strong computational tools to develop numerical solutions.

- Possibility of using numerical and computational solutions is based on a couple of theorems.

- One of such theorems is called the **Law of Large numbers** that is discussed here.

The Law of Large Numbers

- The (weak) law of large numbers states that the sample average converges towards the expected value by increasing the number of samples.

- In mathematical form

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

- In simple words, by increasing the number of sample from rv X, the Mean of samples with probability of 1, will be as adequately close as required to $E[X] = \mu$.

The Law of Large Numbers - Application

- Assume that you wish to calculate $P(X \in A)$ for the continuous rv X.
- This is mathematically equivalent to

$$P(X \in A) = \int_{X \in A} f(x)dx = \int_{X \in A} I_A(x)f(x)dx = E[I_A(x)]$$

in which

$$I_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

- According to the law of large numbers, instead of calculating algebraically, we can use a large number of samples (n) from random variable X and calculate $\overline{I_A(x)}$, as follows, that merges to $E[I_A(x)]$

$$\overline{I_A(x)} = \frac{\sum_i I_A(x_i)}{n} = \frac{\#(I_A(x_i) = 1)}{n}$$

The Law of Large Numbers - Example

Question

If X and Y are independent $U(0,1)$ RVs. Find $P(X + Y < 0.5)$.

Solution

Generate large number of samples for X and Y on the region $U(0,1)$ as shown in the graph by red dots, and see how many percent of them are inside the triangle. This is an estimate for $P(X + Y < 1)$. The larger number of samples, the more accurate answer.

Using R

x <- runif(10000,0,1)

y <- runif(10000,0,1)

I <- which(x+y<0.5)

p <- length(I)/10000

p

0.1264
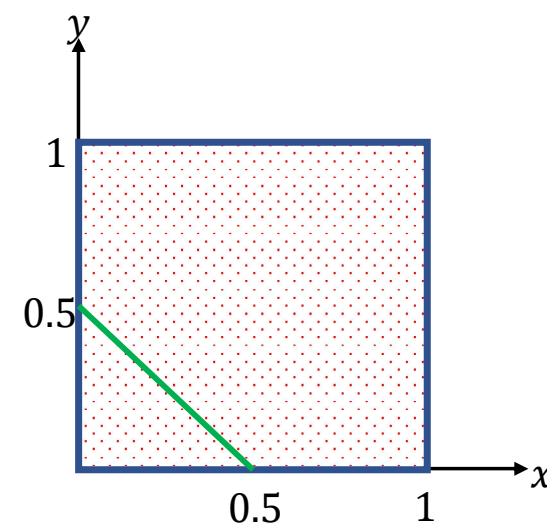
Iterations and Monte-Carlo Method

- By using the simulated random samples as stated in the previous slide, we obviously would not precisely get the same answers for various simulations (iterations), even with the same number of samples.

- In practice, we may therefore be interested in studying the possible variation of the simulated answer.

- To have a more precise answer that removes the effect of randomness in each iteration, we usually find the average of the final answers of all iterations.

- For example, to solve the previous question, we use the method explained in the previous slide using 10,000 samples. The we'll redo the above for 200 iterations to get 200 answers. The average of the answers from 200 iterations is the answer we would rely on.

- Instead of mathematical/algebraic solutions, we can use random samples in repeating algorithms. Such algorithms are generally called Monte-Carlo methods. We should note that Monte-Carlo methods can be used even to answer the questions which may not have random nature.