**CƠ SỞ TOÁN HỌC CHO PHƯƠNG PHÁP MONTE CARLO**

1. **Giới thiệu về Phương pháp MC**

PP MC dựa trên các khái niệm về xác suất và thống kê.

Đây là phương pháp cực kỳ hữu dụng để giải quyết nhiều vấn đề phức tạp

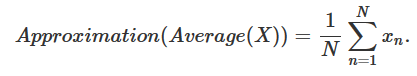
Sử dụng thuật ngữ MC integration.

PP MC có lẽ là phương pháp cũ. Có thể tìm thấy tài liệu đầu tiên trong các ấn phẩm của nhà toán học Comte de Buffon vào đầu thế kỹ thứ 18. Nhưng “Monte Carlo” chỉ được gọi từ những năm 1940. Monte Carlo là tên một quận của vương quốc Monaco. Ở đây nổi tiếng với nhiều song bạc.

Phương pháp Monte Carlo có liên quan đến nhiều lĩnh vực. Nó hữu dụng trong các trò chơi may rủi, những thứ liên quan đến gieo xúc sắc. Những thứ này đều là quá trình ngẫu nhiên. Phương pháp này được các nhà khoa học phát triển như: Fermi, Ulam, von Neumann, Metropolis và những người khác. Và sự phổ biến của nó trong khoa học hiện đại có liên quan nhiều đến máy tính

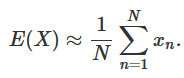
Biến ngẫu nhiên (**random variable**) được ký hiệu bằng chữ in HOA. **(X)**

Trong thống kê được ký hiệu chữ thường (**x1, x2,… là biến ngẫu nhiên của X**)



Trong thống kê, giá trị trung bình của biến ngẫu nhiên X được gọi là **expectation** (kỳ vọng)kí hiệu là **E(X)**

Tích phân MC là một kỹ thuật tính xấp xĩ giá trị kỳ vọng của các biến ngẫu nhiên. Được định nghĩa bằng công thức



1. **Biến ngẫu nhiên và xác xuất.**

Một biến ngẫu nhiên là một hàm X(e) mà khi ánh xạ tập hợp các thí nghiệm sẽ có kết quả là một bộ số. Xác suất (probability) là khái niệm ngắn liền với khái niệm biến ngẫu nhiên.

Quá trình ngẫu nhiên là quá trình không thể dự đoán được kết quả. Tuy nhiên xác xuất thì có thể dự đoán được cơ hôi xảy ra của sự kiện.

Tổng quát hơn, có thể nói rằng nếu kết quả của một quá trình nào đó phải là một trong n kết quả khác nhau và nếu kết quả n này có khả năng xảy ra như nhau, thì xác suất của mỗi kết quả là 1/n.

1. **Phân phối xác suất (part 1)**

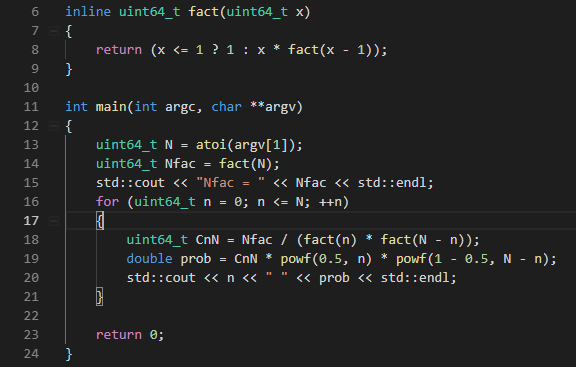
Phân phối xác suất đóng vai trò quan trọng trong phương pháp lấy mẫu trọng (importance sampling)

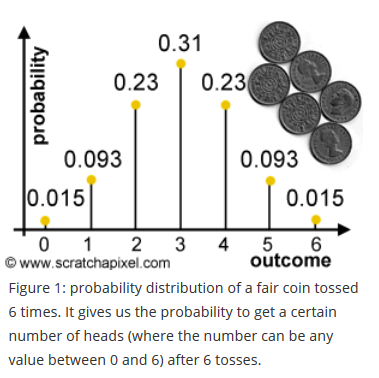
Vì mỗi kết quả có một xác suất liên quan (mà bạn có thể xem là một tập hợp các cặp kết quả và xác suất liên quan của chúng), chúng tôi có thể vẽ các giá trị xác suất này so với các giá trị kết quả có thể có. Trong thống kê, đây là những gì chúng ta gọi là phân phối xác suất.

Trong xác suất và thống kê, phân phối xác suất gán xác suất cho mỗi tập hợp con có thể đo lường được về kết quả có thể có của một thử nghiệm ngẫu nhiên. Hàm phân phối xác suất có thể là rời rạc hoặc liên tục. Xác định xác suất mà biến ngẫu nhiên nhận từng giá trị có thể khác nhau.

**Phân phối nhị thức** (probability distribution)

**Ví dụ:**  Gieo xúc sắc N lần và tính xác suất tất cả các trường hợp xuất hiện n=1,





1. **Tính chất của xác suất**

* For example *A*={1,2,3} and *B*={4,5,6} are two **mutual exclusive** or **disjoint** sets.
* For example, we know that 1, 2, 3, 4, 5, and 6 are the possible outcomes of a die roll. Thus in this case, we can say that the events from that set are **collectively exhaustive** because when we roll a die we are sure to get either one of these numbers (and nothing else than these numbers).
* However if you consider a set made of the outcomes 1 and 6, the events are **mutually exclusive** (because they can't occur at the same time) but they are not collectively exhaustive (because you might get a different number such as 2, 3, 4 or 5, thus you don't have the guarantee that you will get either 1 or 6).
* **multiplication rule**: The chances of getting "heads" after the first toss is 12. Because the events are mutually exclusive, your chances to get "heads" on the second toss are still 50 percent, however the combined probability of getting heads twice in a row is the probability of getting "heads" the first time multiplied by the probability of getting heads the second time that is 14.
* **independent events:** if knowing that some event B has occurred does not change the probability of A, then A and B are said to be independent.

1. **Giới thiệu về thống kê (statistics)**

* Mục tiêu đầu tiên của thống kê là cung cấp thông tin của các biến ngẫu nhiên và phân bố xác suất của chúng.
* Nguyên tắc của việc thu thập dữ liệu thống kê là: ý tưởng là chon ngẫu nhiên trong (we select elements in the population)

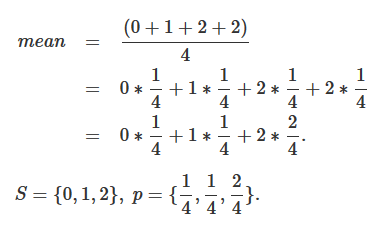
1. **Giá trị kỳ vọng**

* **Giải thích về trọng số (weight):** Ví dụ: S = {0,1,2, 2} ở đây ta thấy số 2 xuất hiền nhiều lần hơn các giá trị còn lại.

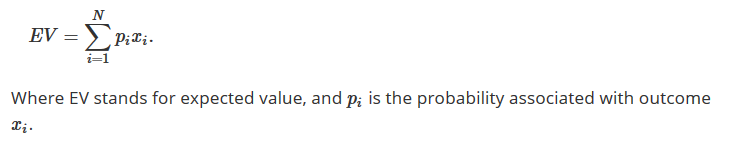
**+** Nghĩa là 2 có trọng số lớn hơn các giá trị còn lại

**+**  Suy ra là **giá trị trung bình** sẽ tiến gần đến số 2.(0+1+2+2)/4=1.25

**Kết luận:** The resulting mean of the card's number in the box is actually the sum of each number multiplied by their individual weight.

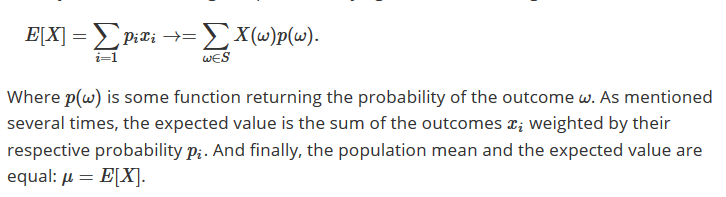


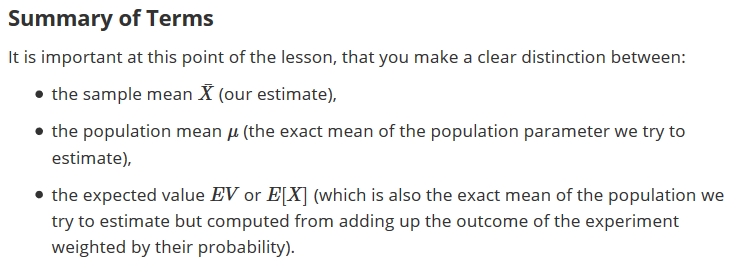
* Giá trị trung bình được tính theo cách này khác so với cách tính giá trị trung bình thông thường. Và phương pháp này được gọi là giá trị kỳ vọng (**expected value**)



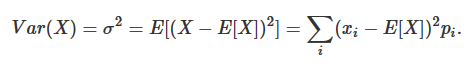
* Nói chung, giá trị trung bình của các biến ngẫu nhiên sẽ hội tụ về một giá trị khi kích thước mẫu tăng. Trong trường giá trị đó như là **giá trị kỳ vọng.**
* **And by definition, this expected value of a random variable is also equal to the**

**IMPORTANT**

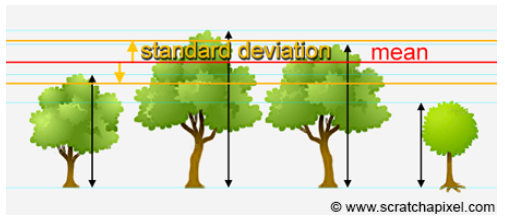




1. Phương sai và độ lệch chuẩn (**variance** and **standard deviation**)







1. Phân phối mẫu (sampling distribution)

**Remind**: what we can do instead is drawing samples from the population (where each of these samples can be seen as a random variable) and dividing the sum of these drawn random variables by the number of samples to get what we call a **sample mean** *X*¯. What this number gives is an **estimation** (not an approximation) of the population **mean *μ*.**

Trong thống kê khi một số đặc tính của quần thể có thể được tính toán sử dụng tất cả các yếu tố hoặc các thành phần trong quần thể. Kết quả đó được gọi là **Parameter** của quần thể. Khi chúng ta sử dụng những mẫu để ước tính (**estimation**) giá trị trung bình của quần thể thì mẫu đó được gọi là thông kê (statistic).

