Các kỹ thuật được phát triển trong luận án này là tất cả các phương pháp Monte Carlo. Phương pháp Monte Carlo là các kỹ thuật số dựa trên việc lấy mẫu ngẫu nhiên để tính gần đúng kết quả của chúng. Tích hợp Monte Carlo áp dụng quy trình này để ước lượng số tích phân. Trong phần phụ lục này, chúng tôi xem xét các khái niệm cơ bản về tích hợp Monte Carlo dựa trên cơ sở của chúng tôi. Từ cuộc thảo luận này, chúng ta sẽ thấy tại sao các phương pháp Monte Carlo là một lựa chọn đặc biệt hấp dẫn cho các vấn đề tích hợp đa chiều phổ biến trong đồ họa máy tính. Tham khảo tốt cho tích hợp Monte Carlo trong bối cảnh đồ họa máy tính bao gồm Pharrand Humphreys [2004], Dutré et al. [2006] và Veach [1997].

Thuật ngữ phương pháp Monte Monte Carlo có nguồn gốc từ Phòng thí nghiệm quốc gia Los Alamos vào cuối những năm 1940 trong quá trình phát triển bom nguyên tử [Metropolis và Ulam, 1949]. Không có gì đáng ngạc nhiên, sự phát triển của các phương pháp này cũng tương ứng với việc phát minh ra các máy tính đầu tiên, giúp tăng tốc đáng kể việc tính toán các nhiệm vụ số lặp đi lặp lại.

Các thuật toán Las Vegas là một loại phương pháp khác dựa trên sự ngẫu nhiên để tính toán các kết quả tính toán. Tuy nhiên, trái ngược với các thuật toán Las Vegas, vốn luôn tạo ra giải pháp chính xác hoặc không chính xác, độ chính xác của các phương pháp Monte Carlo chỉ có thể được phân tích từ một quan điểm thống kê. Vì điều này, trước tiên chúng tôi xem xét một số nguyên tắc cơ bản từ lý thuyết xác suất trước khi mô tả chính xác sự tích hợp Monte Carlo.

**A.1 Probability Background**

Để xác định tích hợp Monte Carlo, chúng tôi bắt đầu bằng cách xem xét một số ý tưởng cơ bản từ xác suất.

**A.1.1 Random Variables**

Một biến ngẫu nhiên X là một hàm ánh xạ kết quả của một quá trình ngẫu nhiên thành các con số. Một biến ngẫu nhiên có thể là rời rạc (ví dụ: một xúc xắc có sáu mặt trong đó có tập giá trị cố định, X = {1,2,3,4,5,6}) hoặc liên tục (ví dụ: một người chiều cao, có thể lấy giá trị thực R.

**A.1.2 Cumulative Distributions and Density Functions (**Hàm phân phối tích lũy và hàm mật độ

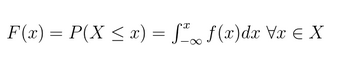
**)**

Theo quy ước, chữ *F* hoa được dùng cho hàm phân phối tích lũy, còn chữ *f* thường được dùng cho [hàm mật độ xác suất](https://vi.wikipedia.org/wiki/H%C3%A0m_m%E1%BA%ADt_%C4%91%E1%BB%99_x%C3%A1c_su%E1%BA%A5t)(PDF cho biến liên tục) và [probability mass function](https://vi.wikipedia.org/w/index.php?title=Probability_mass_function&action=edit&redlink=1) (PMF cho biến rời rạc).

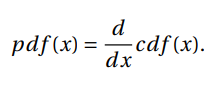
**Hàm mật độ xác suất** (pdf – probability density function) được đặc trưng cho **biến ngẫu nhiên liên tục.** Để tìm xác suất của một biến ngẫu nhiên liên tục, thường ta tính diện tích phần dưới đường cong nằm giữa 2 điểm cần tính xác suất.



Hàm phân phối tích lũy(CDF)của biến ngẫu nhiên X được đặc trưng cho **biến ngẫu nhiên rời rạc và liên tục.** Là hàm F(x) được tính bằng tổng các xác suất của biến ngẫu nhiên X nhỏ hơn hay bằng giá trị x.



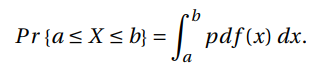
Hàm mật độ xác suất (probability density function:PDF) tương ứng là đạo hàm của CDF:

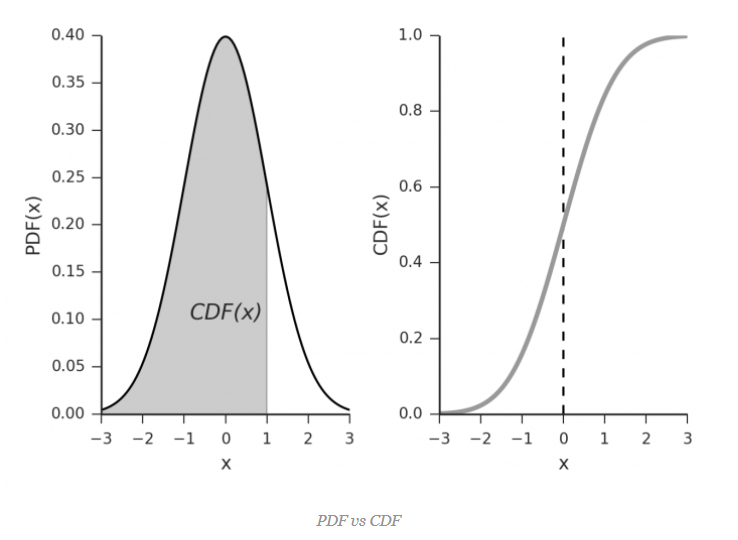


CDF luôn tăng đơn điệu, điều đó có nghĩa là PDF luôn không âm.

Một mối quan hệ quan trọng nảy sinh từ hai phương trình trên, cho phép chúng ta tính toán

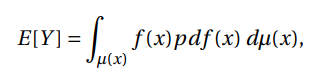
xác suất rằng một biến ngẫu nhiên nằm trong một khoảng:





**A.1.3 Giá trị kỳ vọng và phương sai (Expected Values and Variance)**

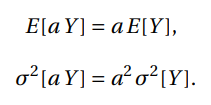
Giá trị kỳ vọng của một biến ngẫu nhiên Y = f (X) trên một miền u(x) được định nghĩa:



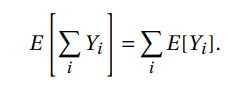
Và phương sai:



trong đó σ, độ lệch chuẩn, là căn bậc hai của phương sai. Từ những định nghĩa này, nó là dễ dàng cho thấy rằng với bất kỳ hằng số a,



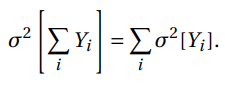
Hơn nữa, giá trị kỳ vọng của một tổng các biến ngẫu nhiên Yi là tổng giá trị kỳ vọng của Yi:



Từ các tính chất này, có thể rút ra biểu thức đơn giản hơn cho phương sai:

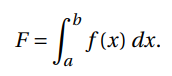


Ngoài ra, nếu các biến ngẫu nhiên không tương quan, thuộc tính tổng cũng giữ cho phương sai

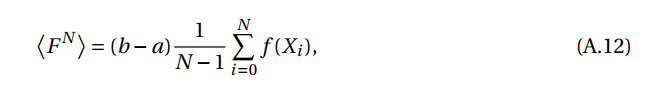


**A.2 Tích phân Monte Carlo**

Tích hợp Monte Carlo sử dụng lấy mẫu ngẫu nhiên của một hàm để ước tính tích phân của nó. Giả sử ta muốn tính tích phân 1D của hàm f(x) từ a đến b:



Chúng ta có thể tính gần đúng tích phân này bằng cách lấy trung bình các mẫu của hàm f tại các điểm ngẫu nhiên đồng nhất trong khoảng. Cho một tập hợp N biến ngẫu nhiên đều Xi € [a, b) với hàm mật độ (PDF) là 1 / (b - a), ước tính Monte Carlo cho F là

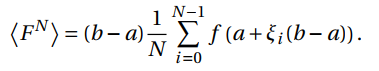


(**A12**)

Biến ngẫu nhiên Xi € [a, b) có thể được xây dựng bằng cách làm cong một số ngẫu nhiên chính tắc được phân bố đồng đều giữa 0 và 1, £i €[0,1):

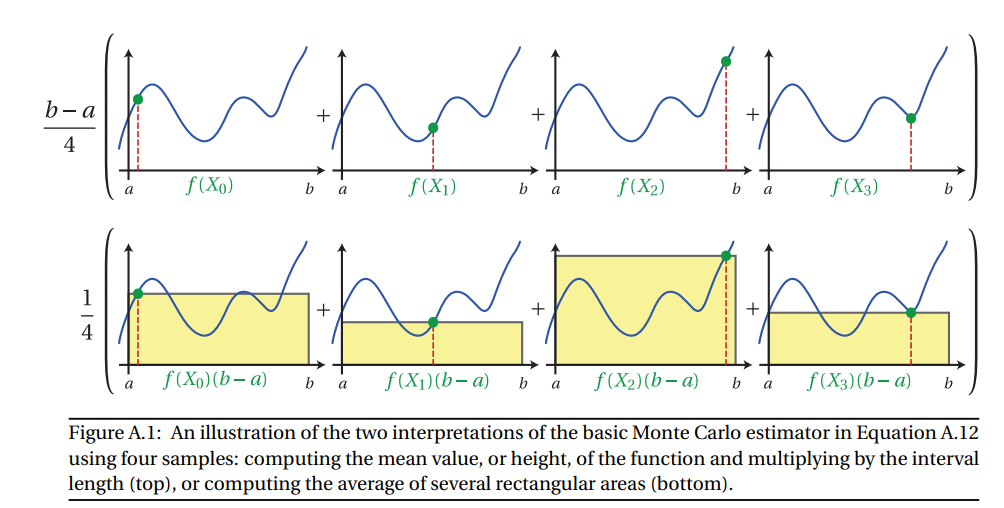


Mở rộng ước tính:



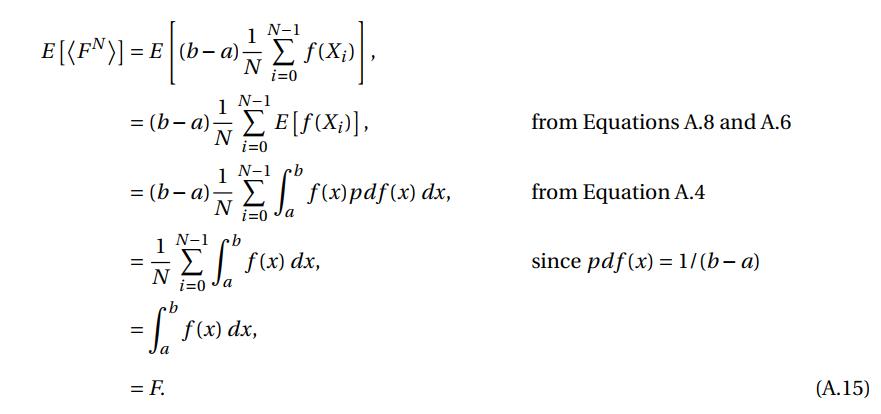
Vì là hàm của Xi, bản thân nó là một biến ngẫu nhiên.Sử dụng ký hiệu này để làm rõ rằng là một xấp xỉ của F khi sử dụng N mẫu.

Nhân thấy rằng tích phân Monte Carlo trong phương trình A.12 tính giá trị trung bình của hàm f(x) trong khoảng từ a đến b, sau đó nhân giá trị trung bình này với độ dài của khoảng (b-a). Bằng cách di chuyển (b-a)vào phép tính tổng, tích phân có thể được tính bằng cách chọn một chiều cao ngẫu nhiên theo giá trị f(Xi)và tính trung bình tập hợp các diện tích hình chữ nhật. Diện tích hình chữ nhật được tính bằng cách nhân chiều cao với chiều dài khoảng (b-a). Được minh họa trong hình A.1



**A.2.1 Giá trị mong đợi và sự hội tụ**

Giá trị kỳ vọng của trên thực tế là:

––

Hơn nữa, khi tăng số lượng mẫu N, ước tính sẽ tiến gần với giá trị F. Do định luật số lượng lớn thì**:**



Trong thực tế, việc xem xét tích phân hội tụ nhanh đến mức nào để có thể đưa ra giải pháp đủ chính xác. Điều này được phân tích bằng cách xác định tốc độ hội tự của phương sai. Và độ lệch chuẩn có tỉ lệ thuận:



Thật không may, điều này có nghĩa là phải tăng gấp bốn lần số lượng mẫu để giảm một nửa sai số!

Các kỹ thuật tính tích phân tiêu chuẩn khác (**tính hinh thang**) sẽ hội tụ nhanh hơn khi tính bằng tích phân MC trong một chiều. Tuy nhiên các kỹ thuật này cùng gặp vấn đề, trong đó tỷ lệ hội tụ trở nên tồi tệ hơn theo cấp số nhân với kích thước chiều tăng lên. Tích phân Monte Carlo cơ bản ở trên có thể dễ dàng được mở rộng ra nhiều chiều. Và tốc độ hội tụ cho Monte Carlo không phụ thuộc vào số lượng kích thước chiều trong tích phân. Điều này làm cho tích phân Monte Carlo trở thành kỹ thuật thực tế duy nhất cho việc giải quyết các bài toán nhiều chiều.

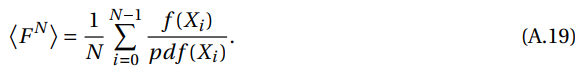
Cách chứng minh vấn đề hội tụ của MC không phụ thuộc số lượng chiều của bài toán:  [https://en.wikipedia.org/wiki/Monte\_Carlo\_integration](%20https:/en.wikipedia.org/wiki/Monte_Carlo_integration)

**A.2.2 Tích phân đa chiều**

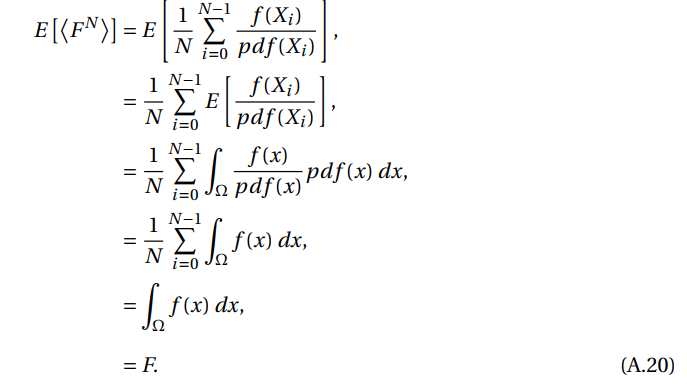
Tích hợp Monte Carlo có thể được tổng quát hóa để sử dụng các biến ngẫu nhiên được rút ra từ các hàm phân phối tích lũy tùy ý. Để tính các tích phân đa chiều, chẳng hạn như



Với sự thay đổi từ phương trình A.12:



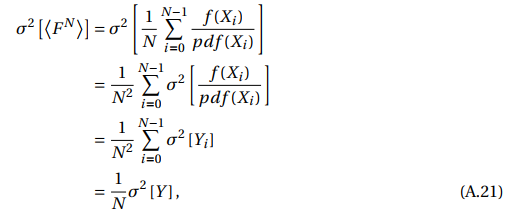
Với giá trị kỳ vọng:



Như vậy ngoài ưu điểm về tốc độ hội tụ không phụ thuộc vào số chiều. Phương pháp tích phân MC còn cho phép tùy ý lựa chọn số lượng mẫu. Như đã đề cập thì tốc độ hội tụ tỉ lệ thuận O(sprt(N)) với số chiều bất kỳ.

A.3 **Variance Reduction (Giảm phương sai)**

Giảm phương sai thì kết quả tích phân sẽ trở nên chính xác hơn. Do không phụ thuộc vào số lượng mẫu và sử dụng phương trình A.10 thì phương sai của <F^N>:



Và độ lệch chuẩn là:



Tại  

Phương trình A.22 đã chứng minh rằng lệch chuẩn hội tụ với O (sqrtN. Biểu thức này cũng cho thấy bằng cách giảm phương sai của mỗi Yi, chúng ta có thể giảm phương sai của 

**A.3.1 Lấy mẫu trọng (Importance Sampling)**

Lấy mẫu trọng làm giảm phương sai bằng cách chọn tùy ý hàm mật độ xác suất (PDF). Nhưng nếu chọn hàm mật độ pdf(x) tương tự với hình dạng của hàm f(x) thì kết quả tích phân càng chính xác. Nhận thấy rằng nếu đặt nhiều mẫu tại những nơi mà tích phân có đóng góp cao. Thì phương sai của tích phân MC giảm đáng kể.