Applied Statistical Mechanics
Lecture Note - 10

Basic Statistics and Monte-Carlo Method -2

고려대학교
화공생명공학과
강정원

Table of Contents

1. General Monte Carlo Method
2. Variance Reduction Techniques
3. Metropolis Monte Carlo Simulation

1.1 Introduction

고려대학교

- Monte Carlo Method
  - Any method that uses random numbers
  - Random sampling the population
  - Application
    - Science and engineering
    - Management and finance
- For given subject, various techniques and error analysis will be presented
- Subject : evaluation of definite integral

$$I = \int_a^b \rho(x)dx$$

1.1 Introduction

- Monte Carlo method can be used to compute integral of any dimension ***d*** (***d***-fold integrals)
- Error comparison of d-fold integrals
  - Simpson's rule,… $E \propto N^{-1/d}$
  - Monte Carlo method $E \propto N^{-1/2}$ → purely statistical, not rely on the dimension !
  - Monte Carlo method WINS, when d >> 3

1.2 Hit-or-Miss Method

- Evaluation of a definite integral

$$I = \int_a^b \rho(x)dx$$

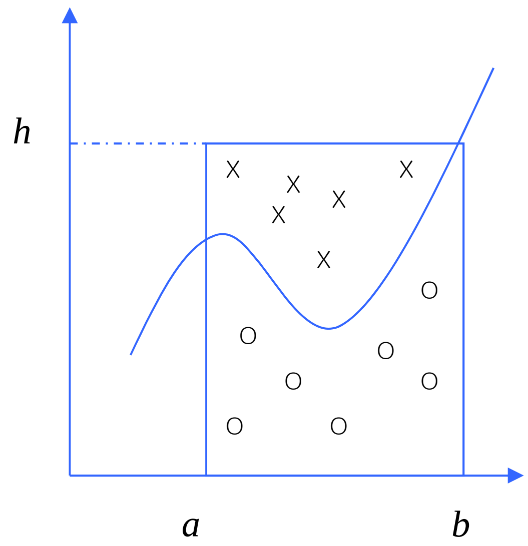
$$h \geq \rho(x) \text{ for any } x$$

- Probability that a random point reside inside the area

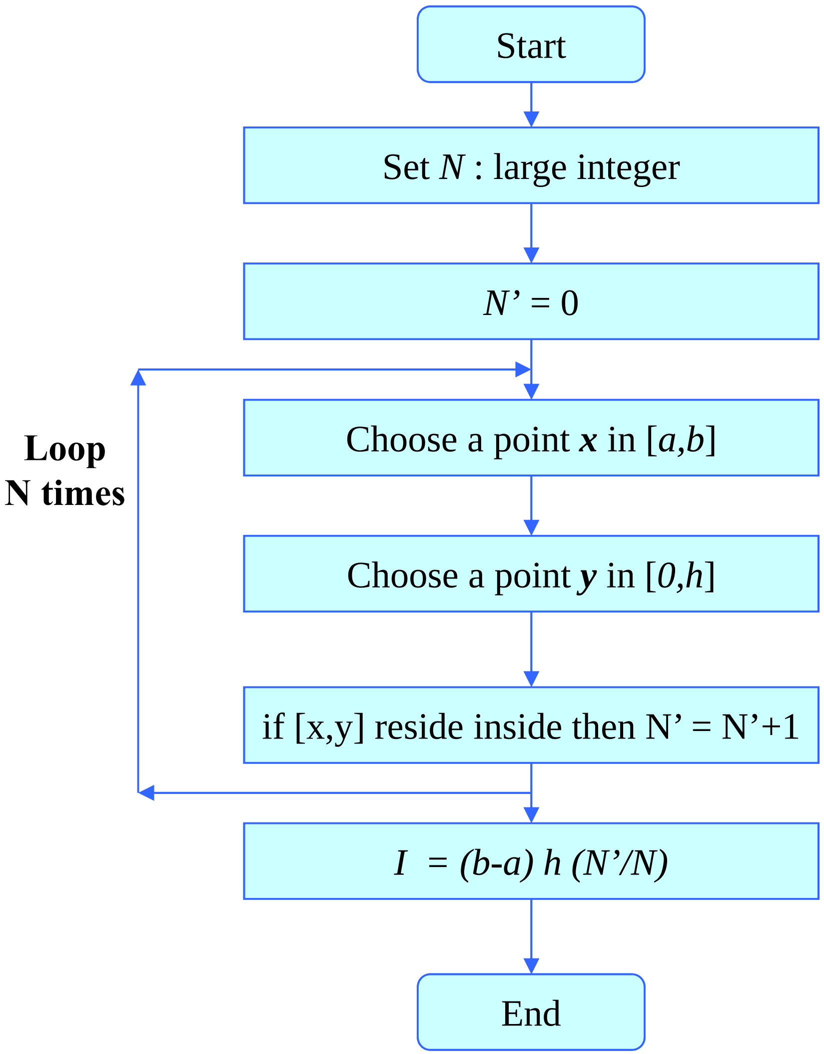
$$r = \frac{I}{(b-a)h} \approx \frac{N'}{N}$$

  - N : Total number of points
  - N' : points that reside inside the region

$$I \approx (b-a)h\frac{N'}{N}$$

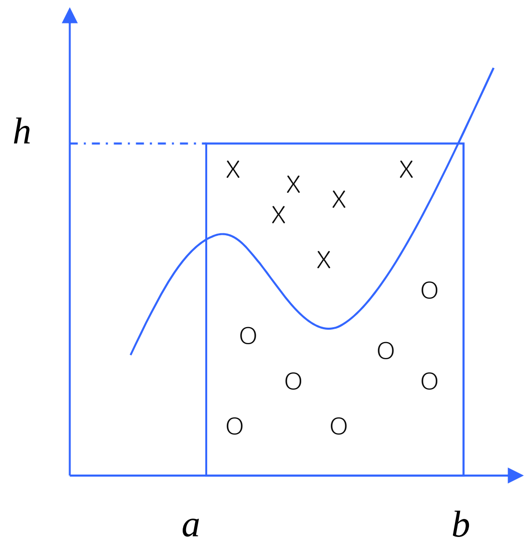


1.2 Hit-or-Miss Method



$$x = (b-a)u_1 + a$$

$$y = hu_2$$



1.2 Hit-or-Miss Method

- Error Analysis of the Hit-or-Miss Method
  - It is important to know how accurate the result of simulations are
  - The rule of 3σ's
- Identifying Random Variable

$$\overline{X} = \frac{1}{N}\sum_{n=1}^{N} X_n$$

  - From central mean theorem , $\overline{X}$ is normal variable in the limit of large N

each X_n		$\overline{X}$
mean value: μ	⇨	mean value: μ
variance : σ^2		variance : σ^2 / N

1.2 Hit-or-Miss Method

고려대학교

- Sample Mean : estimation of actual mean value (μ)

$$\mu' = \frac{1}{N}\sum_{n=1}^{N} x_n$$

$$\mu' = \frac{1}{N}\sum_{n=1}^{N} x_n = r = \frac{N'}{N}$$

$$\sum_{n=1}^{N} x_n = N'$$

x	0	1
p(x)	r	1-r

- Accuracy of simulation → the most probable error

$$\frac{0.6745\sigma}{\sqrt{N}} \qquad \sigma = \sqrt{V(X)}$$

1.2 Hit-or-Miss Method

- Estimation of error

$$\frac{0.6745\sigma}{\sqrt{N}} \qquad \sigma = \sqrt{V(X)}$$

- We do not know exact value of s , m
- We can also estimate the variance and the mean value from samples …

$$\sigma'^2 = \frac{1}{N-1}\sum_{n=1}^{N}(x_n - \mu')^2$$

1.2 Hit-or-Miss Method

- For present problem (evaluation of integral) exact answer (I) is known → estimation of error is,

$$V(X) = E(X^2) - [E(X)]^2$$

$$E(X) = \sum xp(x) = 1\times r + 0\times(1-r) = r = \mu$$
$$E(X^2) = \sum x^2 p(x) = 1\times r + 0\times(1-r) = r$$

$$V(X) = r - r^2 = (1-r)r = \sigma^2$$
$$\sigma = \sqrt{(1-r)r}$$

$$I = r(a-b)h = \mu(a-b)h$$

$$I_{error}^{HM} = \frac{0.6745\sigma}{\sqrt{N}} = 0.6745\times(b-a)h\sqrt{\frac{r(1-r)}{N}} = 0.6745\sqrt{\frac{I((b-a)h-I)}{N}}$$

1.3 Sample Mean Method

- $\rho(x)$ is a continuous function in x and has a mean value ;

$$< \rho >= \frac{\int_a^b \rho(x)dx}{\int_a^b dx} = \frac{1}{b-a}\int_a^b \rho(x)dx$$

$$\therefore I = \int_a^b \rho(x)dx = (b-a) < \rho >$$

$$< \rho >= \frac{1}{N}\sum_{n=1}^{N} \rho(x_n)$$

$$x_n = (b-a)u_n + a \qquad u_n = \text{uniform random variable in } [0,1]$$

1.3 Sample Mean Method

- Error Analysis of Sample Mean Method
  - Identifying random variable

$$\overline{Y} = \frac{1}{N}\sum_{n=1}^{N} Y_n$$

$$\mu_{\overline{Y}} = <\rho> \approx \frac{1}{N}\sum_{n=1}^{N} y_n = \frac{1}{N}\sum_{n=1}^{N} \rho(x_n)$$

  - Variance

$$<\rho^2> = \frac{1}{N}\sum_{n=1}^{N} \rho(x_n)^2$$

$$V(\rho) = <\rho^2> - [<\rho>]^2$$

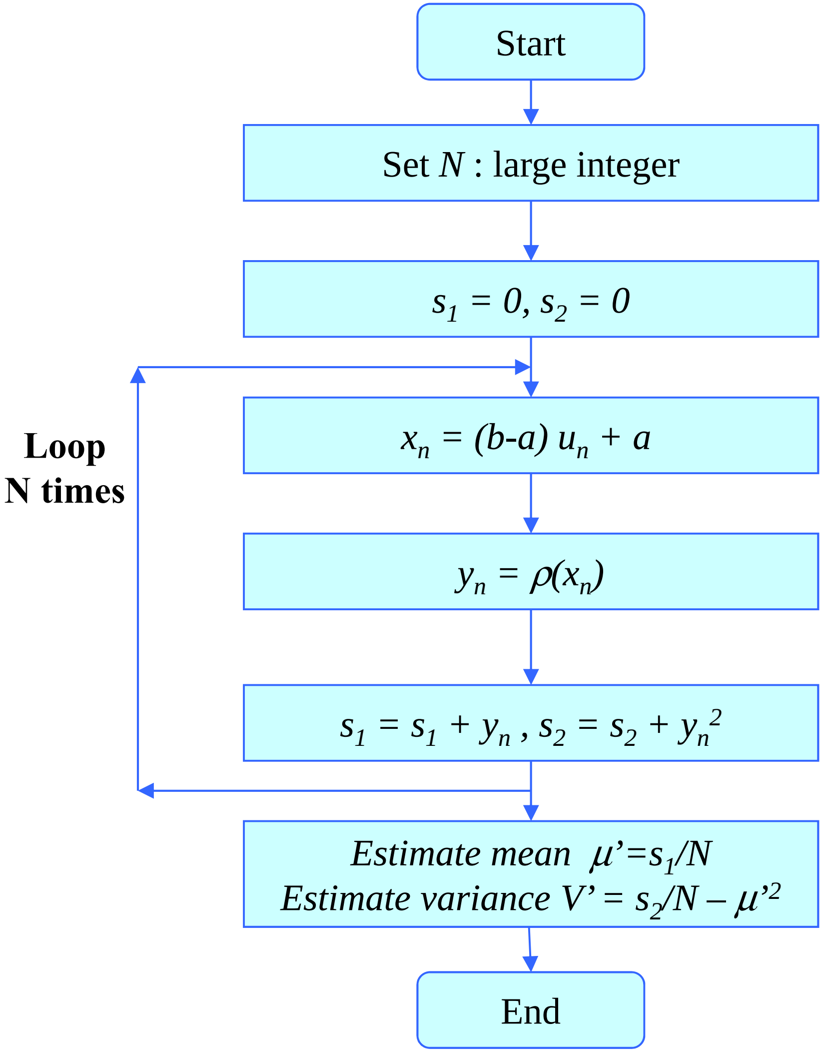
$$I_{error}^{SM} = 0.6745\sigma = (b-a)\times 0.6745\sqrt{\frac{V(x)}{N}}$$

1.3 Sample Mean Method

- If we know the exact answer,

$$I_{error}^{SM} = 0.6745\sqrt{\frac{L\int_a^b \rho(x)^2 dx - I^2}{N}}$$

1.3 Sample Mean Method



$$I_{error}^{SM} = (b-a) \times 0.6745 \sqrt{\frac{V'}{N}}$$

QUIZ

- Compare the error for the integral

$$I = \int_0^{\pi/2} \sin x dx = -\cos x\Big|_0^{\pi/2} = 1$$

using HM and SM method

$$I_{error}^{HM} = \frac{0.6745\sigma}{\sqrt{N}} = 0.6745 \times (b-a)h\sqrt{\frac{r(1-r)}{N}} = 0.6745\sqrt{\frac{I((b-a)h-I)}{N}}$$

$$I_{error}^{SM} = 0.6745\sqrt{\frac{L\int_a^b \rho(x)^2 dx - I^2}{N}}$$

Example : Comparison of HM and SM

고려대학교

- Evaluate the integral

$$I = \int_0^{\pi/2} \sin x dx = -\cos x \Big|_0^{\pi/2} = 1$$

$$I_{error}^{HM} = 0.6745\sqrt{\frac{I((b-a)h-I)}{N}} = 0.6745\sqrt{\frac{1((\pi/2-0)\times 1-1)}{N}} = 0.6745\sqrt{\frac{\pi/2-1}{N}}$$

$$<\rho>= I/(b-a) = 2/\pi$$

$$<\rho^2>= \frac{1}{b-a}\int_0^{\pi/2} \sin^2 x dx = 1/2$$

$$V(\rho) = \frac{1}{2} - \left(\frac{2}{\pi^2}\right)$$

$$I_{error}^{SM} = 0.6745(b-a)\sqrt{\frac{V'}{N}} = 0.6745\sqrt{\frac{(\pi^2/4)(1/2-4/\pi^2)}{N}} = 0.6745\sqrt{\frac{\pi^2/8-1}{N}}$$

Example : Comparison of HM and SM

- Comparison of error

$$\frac{I_{error}^{SM}}{I_{error}^{HM}} = \sqrt{\frac{\frac{\pi^2}{8}-1}{\frac{\pi}{2}-1}} \approx 0.64$$

SM method has 36 % less error than HM

- No of evaluation having the same error

$$N^{SM} = \frac{\frac{\pi^2}{8}-1}{\frac{\pi}{2}-1} N^{HM} \approx 0.41 N^{HM}$$

SM method is more than twice faster than HM

2.1 Variance Reduction Technique - Introduction

- Monte Carlo Method and Sampling Distribution
  - Monte Carlo Method : Take values from random sample
  - From central limit theorem,

$$\overline{\mu} = \mu \qquad\qquad \overline{\sigma}^2 = \sigma^2 / N$$

  - 3s rule

$$P(\mu - 3\overline{\sigma} \le \overline{X} \le \mu - 3\overline{\sigma}) \approx 0.9973$$

  - Most probable error

$$Error \approx \pm \frac{0.6745\sigma}{\sqrt{N}}$$

- Important characteristics

$$Error \propto 1/\sqrt{N} \qquad\qquad Error \propto \sigma$$

2.1 Variance Reduction Technique - Introduction

- Reducing error
  - *100 samples reduces the error order of 10
  - Reducing variance → Variance Reduction Technique

- The value of variance is closely related to how samples are taken
  - Unbiased sampling
  - Biased sampling
    - More points are taken in important parts of the population

2.2 Motivation of Variance Reduction Technique

- If we are using sample-mean Monte Carlo Method
  - Variance depends very much on the behavior of $\rho(x)$

    $\rho(x)$ varies little → variance is small

    $\rho(x)$ = const → variance=0
- Evaluation of a integral

$$I' = (b-a)\mu_{\overline{Y}} = \frac{b-a}{N}\sum_{n=1}^{N}\rho(x_n)$$

  - Near minimum points → contribute less to the summation
  - Near maximum points → contribute more to the summation

  - More points are sampled near the peak →"importance sampling strategy"

2.3 Variance Reduction using Rejection Technique

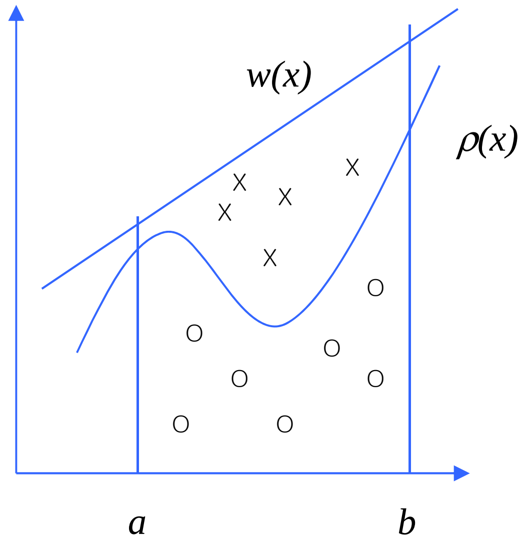
- Variance Reduction for Hit-or-Miss method
- In the domain [a,b] choose a comparison function

$$w(x) \geq \rho(x)$$

$$W(x) = \int_{-\infty}^{x} w(x)dx$$

$$A = \int_{a}^{b} w(x)dx$$

$$Au = w(x) \longrightarrow x = W^{-1}(Au)$$



- Points are generated on the area under *w(x)* function
  - Random variable that follows distribution *w(x)*

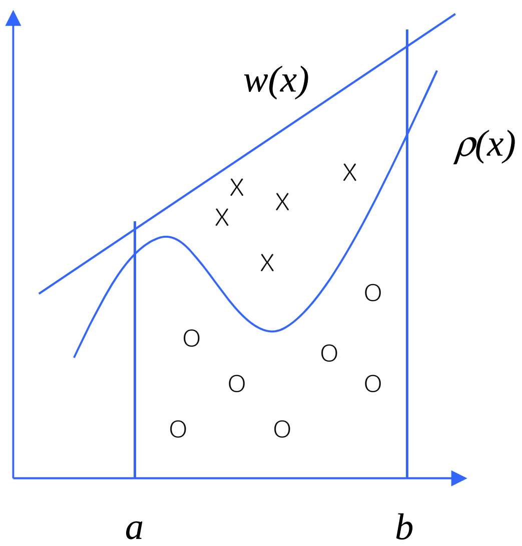
2.3 Variance Reduction using Rejection Technique

- Points lying above $\rho(x)$ is rejected

$$I \approx A\frac{N'}{N}$$

$$y_n = w(x_n)u_n$$

$$q_n = \begin{cases} 1 \text{ if } y_n \le \rho(x_n) \\ 0 \text{ if } y_n > \rho(x_n) \end{cases}$$



q	1	0
P(q)	r	1-r

$$r = I / A$$

2.3 Variance Reduction using Rejection Technique

고려대학교

- Error Analysis

$$E(Q) = r, \qquad V(Q) = r(1-r)$$

$$I = Ar = AE(Q)$$

$$I_{error}^{RJ} \approx 0.67\sqrt{\frac{I(A-I)}{N}}$$

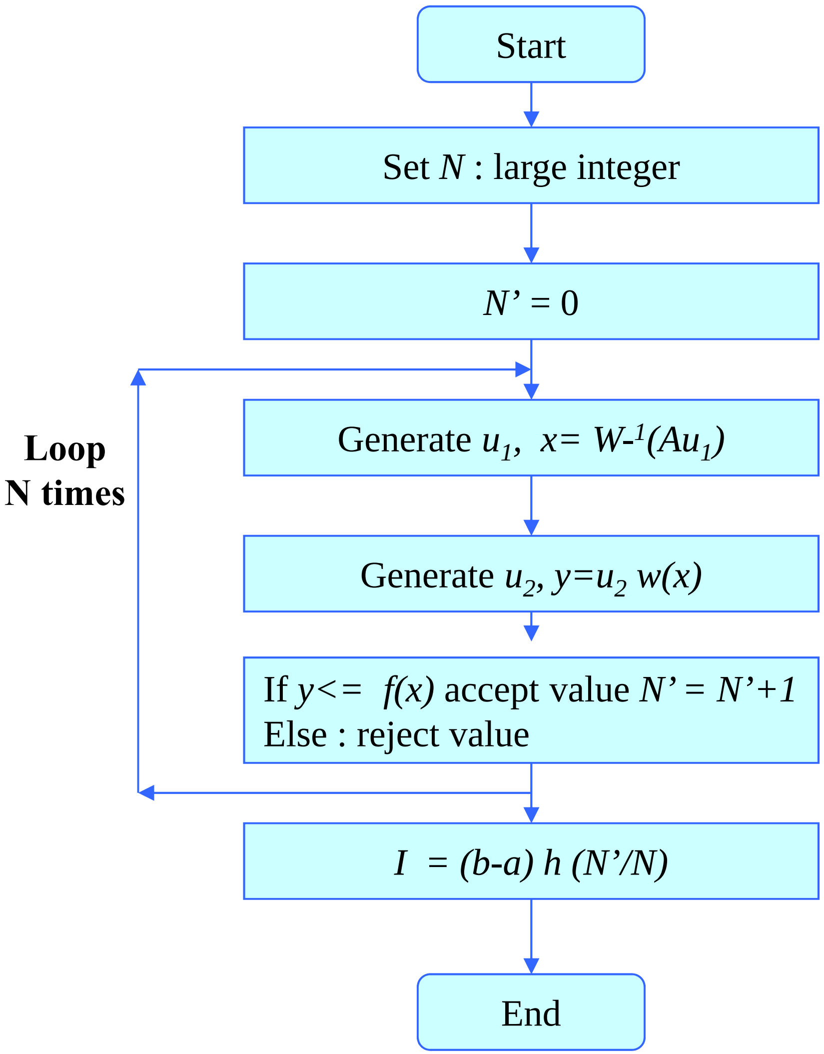
Hit or Miss method

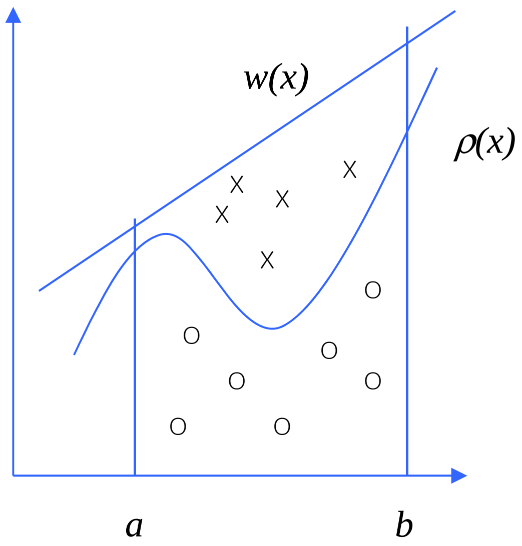
$$A = (b-a)h$$

Error reduction

$$A \to I \quad \text{then} \quad \text{Error} \to 0$$

2.3 Variance Reduction using Rejection Technique





$$I_{error}^{RJ} \approx 0.67\sqrt{\frac{I'(A-I')}{N}}$$

2.4 Importance Sampling Method

- Basic idea
  - Put more points near maximum
  - Put less points near minimum

- F(x) : transformation function (or weight function_

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

$$y = F(x)$$

$$x = F^{-1}(y)$$

2.4 Importance Sampling Method

고려대학교

$$dy/dx = f(x) \rightarrow dx = dy/f(x)$$

$$I = \int_a^b \frac{\rho(x)}{f(x)} dy = \int_a^b \left[\frac{\rho(x)}{f(x)} \right] f(x) dx$$

$$\gamma(x) = \frac{\rho(x)}{f(x)}$$

$$< \eta >_f = \int_a^b \eta(x) f(x) dx$$

if we choose $f(x) = c\rho(x)$,then variance will be small
The magnitude of error depends on the choice of $f(x)$

$$I = \int_a^b \gamma(x) f(x) dx = < \gamma >_f$$

2.4 Importance Sampling Method

- Estimate of error

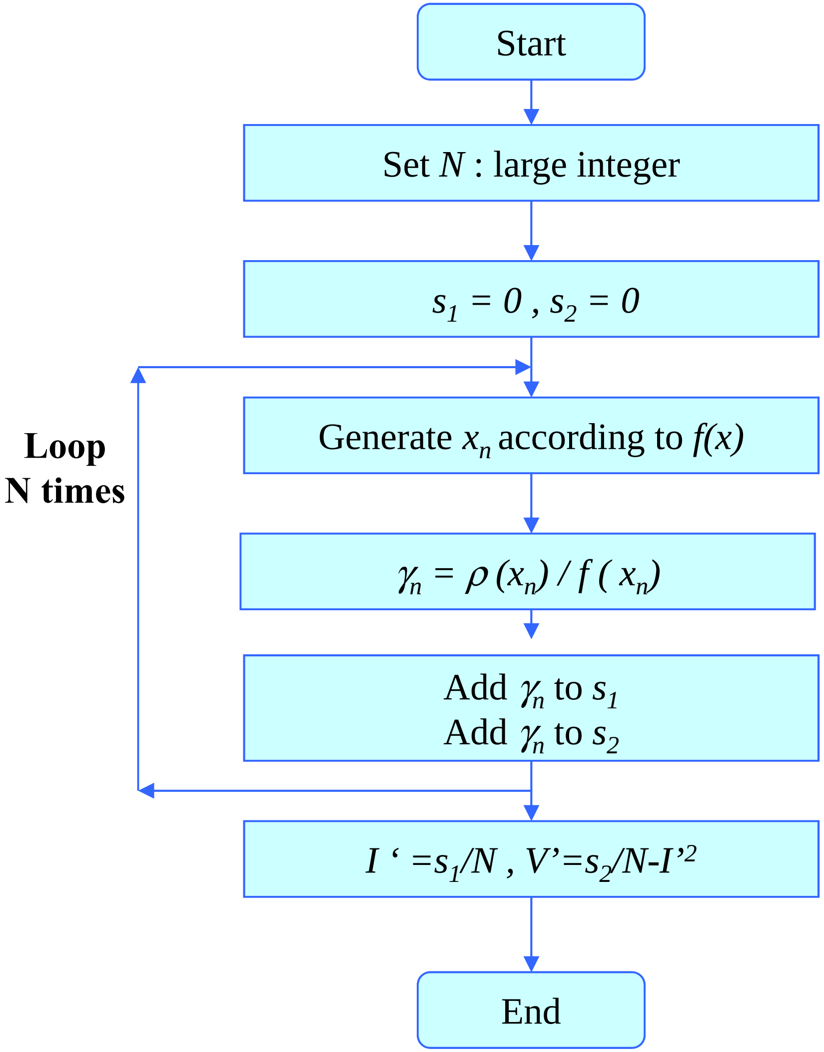
$$I = \langle \gamma \rangle_f \approx \frac{1}{N}\sum_{n=1}^{N}\gamma(x_n)$$

$$I_{error} = 0.67\sqrt{\frac{V_f(\gamma)}{N}}$$

$$V_f(\gamma) = \langle \gamma^2 \rangle_f - (\langle \gamma \rangle_f)^2$$

$$I_{error}^{IS} = 0.67\sqrt{\frac{\langle \gamma^2 \rangle_f - I^2}{N}}$$

2.4 Importance Sampling Method



$$I^{IS}_{error} = 0.67\sqrt{\frac{V'}{N}}$$

3. Metropolis Monte Carlo Method and Importance Sampling

고려대학교

- Average of a property in Canonical Ensemble

$$< A >_{NVT} = \frac{\int A(\mathbf{r}^N)\exp(-U(\mathbf{r}^N)/kT)d\mathbf{r}^N}{\int \exp(-U(\mathbf{r}^N)/kT)d\mathbf{r}^N}$$

$$= \int \frac{\exp(-U(\mathbf{r}^N)/kT)}{Z(\mathbf{r}^N)} A(\mathbf{r}^N)d\mathbf{r}^N$$

$$= \int \mathcal{N}(\mathbf{r}^N)A(\mathbf{r}^N)d\mathbf{r}^N$$

Probability

3. Metropolis Monte Carlo Method and Importance Sampling

- Create n_i random points in a volume $\mathbf{r}_i^N$ such that

$$n_i = \mathcal{N}(\mathbf{r}_i^N) L$$

$$< A >_{NVT} = \frac{1}{L} \sum_{trials} n_i A(\mathbf{r}^N)$$

- Problem : How we can generate n_i random points according to $\mathcal{N}(\mathbf{r}_i^N)$?

→ We cannot use inversion method

→ Use Markov chain with Metropolis algorithm

3. Metropolis Monte Carlo Method and Importance Sampling

- Markov chain :Sequence of stochastic trials satisfies few some conditions
  - Stochastic process that has no memory
  - Selection of the next state only depends on current state, and not on prior state
  - Process is fully defined by a set of transition probabilities π_{ij}

π_{ij} = probability of selecting state *j* next, given that presently in state *i*.

Transition-probability matrix Π collects all π_{ij}

Markov Chain

- Notation
  - Outcome $\boldsymbol{\rho}$
  - Transition matrix $\boldsymbol{\pi}$
- Example
  - Reliability of a computer
    - if it is running 60 % of running correctly on the next day
    - if it is down it has 80 % of down on the next day

$$\boldsymbol{\pi} = \begin{pmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{pmatrix}$$

Markov Chain

$$\boldsymbol{\rho}^{(1)} = (0.5 \quad 0.5)$$

$$\boldsymbol{\rho}^{(2)} = \boldsymbol{\rho}^{(1)}\boldsymbol{\pi} = (0.45 \quad 0.55)$$

$$\boldsymbol{\rho}^{(3)} = \boldsymbol{\rho}^{(2)}\boldsymbol{\pi} = \boldsymbol{\rho}^{(1)}\boldsymbol{\pi}^2 = (0.435 \quad 0.565)$$

$$\boldsymbol{\rho} = \lim_{\tau\to\infty} \boldsymbol{\rho}^{(1)}\boldsymbol{\pi}^{\tau} = (0.4286 \quad 0.5714)$$ → limiting behavior always converges to a certain value independent of initial condition

$$\boldsymbol{\rho}\boldsymbol{\pi} = \boldsymbol{\rho} \quad \longrightarrow \quad \sum_m \rho_m \pi_{mn} = \rho_n$$

$$\sum_n \pi_{mn} = 1$$ → Stochastic matrix : sum of the probability should be 1

Features
- Every state can be eventually reached from another state
- The resulting behavior follows a certain probability