CS303: Mathematical Foundations for AI

Sampling

28 Mar 2025

Recap

- Previous

Recap

- Previous
  - Change of Variables (Discrete)

Recap

- Previous
  - Change of Variables (Discrete)
  - Change of Variables (Continuous)

Recap

- Previous
  - Change of Variables (Discrete)
  - Change of Variables (Continuous)
- Inverse CDF Sampling/Inverse transform Sampling/Smirnov Sampling/

Recap

- Previous
  - Change of Variables (Discrete)
  - Change of Variables (Continuous)
- Inverse CDF Sampling/Inverse transform Sampling/Smirnov Sampling/
- Rejection Sampling

Recap

- Previous
  - Change of Variables (Discrete)
  - Change of Variables (Continuous)
- Inverse CDF Sampling/Inverse transform Sampling/Smirnov Sampling/
- Rejection Sampling
- Monte Carlo Sampling

References

- Murphy, Kevin P. Machine learning: a probabilistic perspective. MIT press, 2012. (link)
- Bishop, Christopher M., and Nasser M. Nasrabadi. Pattern recognition and machine learning.(link)
- Sampling (Resource 1) (Video Lectures)

Sampling Uniform Distribution

How to generate $X \sim Unif[0,1]$?

Sampling Uniform Distribution

How to generate $X \sim Unif[0,1]$?

Pseudo Random Number Generators (PRNG)

$$X_{n+1} = (aX_n + c) \mod m$$

Mersenne Twister

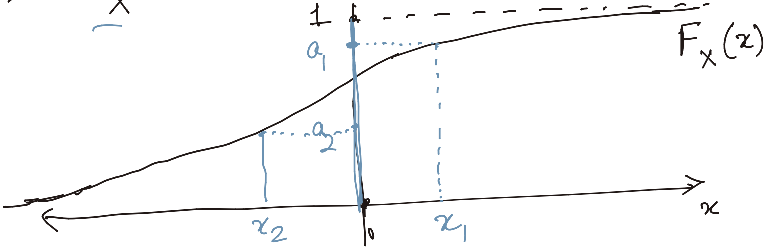
Sampling Non-Uniform Distribution

How to sample any distribution assuming we can sample from Uniform Distribution?

Inverse transform sampling

$X \in \mathbb{R}$

$X \sim f_X$



Inverting the CDF!

$\longrightarrow a_i \quad \text{Unif}[0,1]$

$x \text{ s.t. } F(x) = a$

$\longrightarrow x_1$

$\longrightarrow x_2$

$\vdots$

$\boxed{F^{-1}(a)} \sim f_X$

Inverse transform sampling

Case 1: CDF is invertible

Inverse transform sampling

Case 1: CDF is invertible

- Consider a random variable $X : \Omega \to \mathbb{R}$

Inverse transform sampling

Case 1: CDF is invertible

- Consider a random variable $X : \Omega \to \mathbb{R}$
- F_X is invertible i.e.,

$$\exists\ F^{-1} : (0,1) \to \mathbb{R} \text{ s.t. } F^{-1}\left(F(x)\right) = x,\ \forall x \in \mathbb{R}$$

Inverse transform sampling

Case 1: CDF is invertible

- Consider a random variable $X : \Omega \to \mathbb{R}$
- F_X is invertible i.e.,

$$\exists\ F^{-1} : (0,1) \to \mathbb{R} \text{ s.t. } F^{-1}\left(F(x)\right) = x,\ \forall x \in \mathbb{R}$$

Inverse transform sampling

Case 1: CDF is invertible

- Consider a random variable $X : \Omega \to \mathbb{R}$
- F_X is invertible i.e.,

$$\exists\, F^{-1} : (0,1) \to \mathbb{R} \text{ s.t. } F^{-1}\left(F(x)\right) = x,\ \forall x \in \mathbb{R}$$

Theorem 5

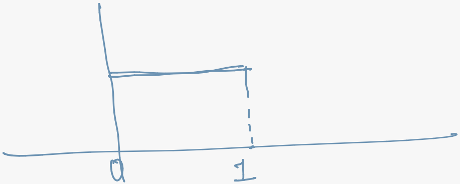
If $U \sim Unif(0,1)$ then $F^{-1}(U) \sim F$ i.e., $F^{-1}(U)$ has a CDF F. PDF f

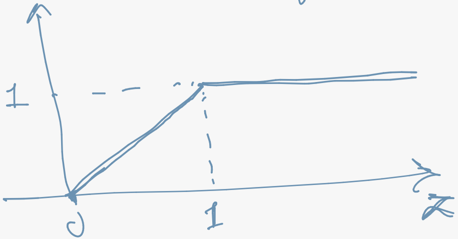
$$\Pr\big(\underbrace{F^{-1}(U)}\leq x\big) = \Pr\big(U \leq \underbrace{\underbrace{F(x)}}_{z}\big)$$

$$= \underline{F(x)}$$

$U \sim \mathrm{Unif}\,\underline{[0,1)}$

$$\Pr(U \leq z) = \int_0^z dz = z = \begin{cases} 0 & z \leq 0 \\ z & 0 < z < 1 \\ 1 & z \geq 1 \end{cases}$$





Invertible CDF

① $\underline{u} \sim Unif[0,1]$

$$X \sim Unif(a,b) \quad F(x) = \int_a^x \frac{1}{b-a}\,dx$$

- $X \sim Unif(a,b)$ (viz)

$$= \frac{x-a}{b-a}$$

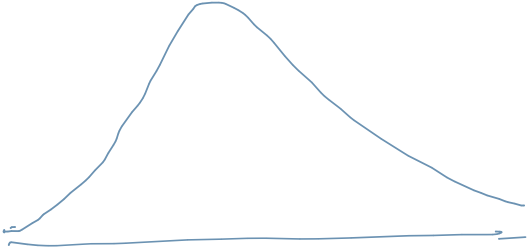
$$F(x) = \frac{x-a}{b-a} = \underline{u}$$

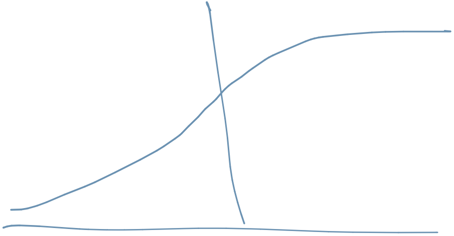
$$\boxed{F^{-1}(u) = a + \underbrace{(b-a)\,u}}$$

Invertible CDF

$$\frac{1}{\sqrt{2\pi\sigma}}\int_0^x e^{-(x-\mu)^2}\,dx$$

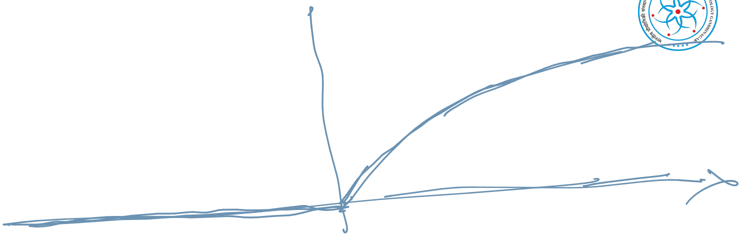
- $X \sim Unif(a,b)$ (viz)
- $X \sim Normal(\mu,\sigma)$?





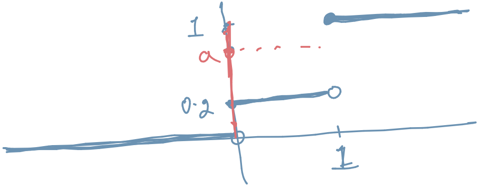
CDF

$\lambda e^{-\lambda x}$



- Exponential
- Discrete

$p(1) = 0.8$



Inverse Transform Sampling

Case 2: CDF is not invertible

Inverse Transform Sampling

Case 2: CDF is not invertible

- Consider a random variable $X : \Omega \to \mathbb{R}$

Inverse Transform Sampling

$x \in [1, \infty)$, $F(x) = 1 \geq 0.5$

$\inf([1, \infty)) = 1$

Case 2: CDF is not invertible

- Consider a random variable $X : \Omega \to \mathbb{R}$
- $G : (0,1) \to \mathbb{R}$

$a = 0.1$

0.1 0.2

$G(0.1)$

$$G(a) = inf\{x \in \mathbb{R} : a \leq F(x)\}$$

$= \inf\{x \in \mathbb{R} : 0.1 \leq F(x)\}$

$= \inf [0, \infty)$

$= 0$

$\{$

Bern.

0

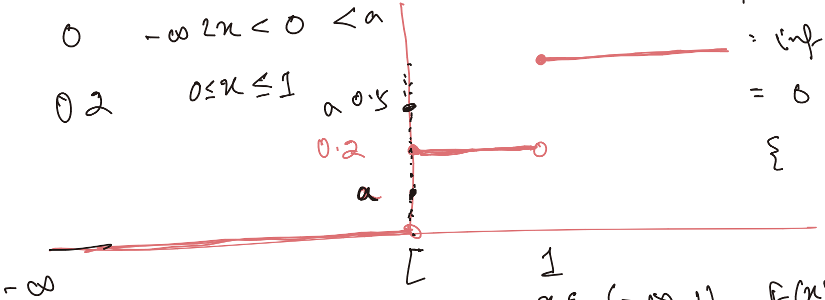
1

0

$-\infty \leq x < 0 < a$

0 2

$0 \leq x \leq 1$



$x \in (-\infty, 1)$ $F(x) \leq 0.2 < 0.5$

Inverse Transform Sampling

Case 2: CDF is not invertible

- Consider a random variable $X : \Omega \to \mathbb{R}$
- $G : (0,1) \to \mathbb{R}$

$$G(a) = inf\{x \in \mathbb{R} : a \leq F(x)\}$$

Inverse Transform Sampling

Case 2: CDF is not invertible

- Consider a random variable $X : \Omega \to \mathbb{R}$
- $G : (0,1) \to \mathbb{R}$

$$G(a) = inf\{x \in \mathbb{R} : a \leq F(x)\}$$

Theorem 10

If $U \sim Unif(0,1)$ then $G(U) \sim F$ i.e., $G(U)$ has a CDF F.

$$\boxed{P(G(U) \leq z) = P(U \leq F(z))}$$

$$= F(z)$$

Inverse CDF

Examples

- Discrete

x_i	1	2	3	4
p_i	0.1	0.4	0.2	0.3

Inverse CDF

Examples

- Discrete

x_i	1	2	3	4
p_i	0.1	0.4	0.2	0.3

  - Generate $U \sim Unif(0,1)$

Inverse CDF

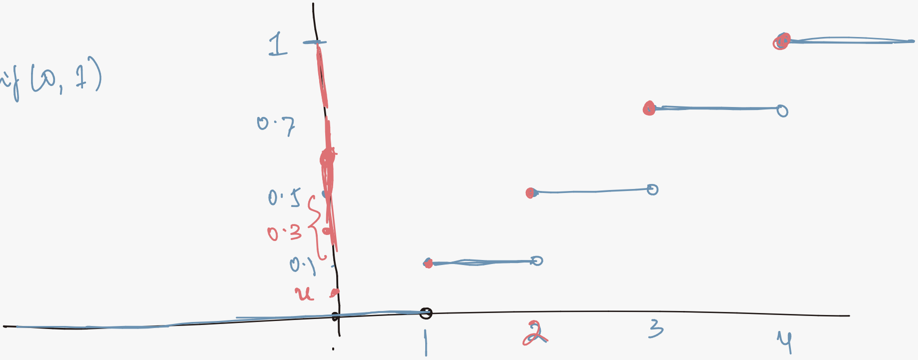
Examples

- Discrete

x_i	1	2	3	4
p_i	0.1	0.4	0.2	0.3

  - Generate $U \sim Unif(0,1)$
  - Determine k s.t. $\sum_{j=1}^{k-1} p_j \le U < \sum_{j=1}^{k} p_j$, return $X = x_k$

$u \sim \text{Unif}(0, 1)$



$$\sum_{j=1}^{k-1} p_j \le u < \sum_{j=1}^{k} p_j$$

$k = 1 \quad : \quad \sum_{j=1}^{k} p_j = 0.1$

$$\sum_{j=1}^{1} p_j = 0.1 \le 0.3 < \sum_{j=1}^{\boxed{2}} p_j = 0.5$$

$k = 2$

$$u < \sum_{j=1}^{1} p_j$$

$k = 1$

Inverse CDF

Examples

- Discrete

x_i	1	2	3	4
p_i	0.1	0.4	0.2	0.3

  - Generate $U \sim Unif(0,1)$
  - Determine k s.t. $\sum_{j=1}^{k-1} p_j \leq U < \sum_{j=1}^{k} p_j$, return $X = x_k$
  - viz

Inverse CDF

Examples

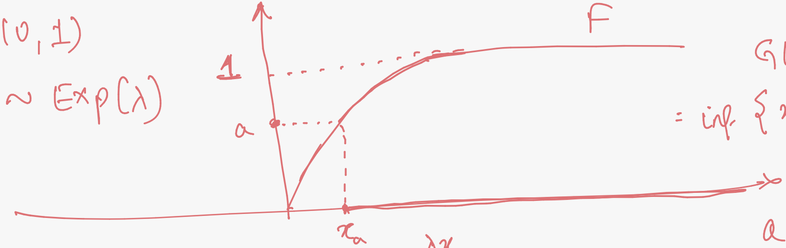
- Discrete

x_i	1	2	3	4
p_i	0.1	0.4	0.2	0.3

  - Generate $U \sim Unif(0,1)$
  - Determine k s.t. $\sum_{j=1}^{k-1} p_j \leq U < \sum_{j=1}^{k} p_j$, return $X = x_k$
  - viz
- $X \sim \lambda e^{-\lambda x}$ (Exponential)

$a \sim \text{unif}(0,1)$

$G(a) \sim \text{Exp}(\lambda)$



$G(a) = \inf\{x : a \le F(x)\}$

$a = 1 - e^{-\lambda x_a}$

$x_a = -\frac{1}{\lambda}\ln(1-a)$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

$$\underline{F_X(x)} = \begin{cases} \underline{1 - e^{-\lambda x}} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

$$\boxed{G(a) = -\frac{1}{\lambda}\ln(1-a)}$$

Box-Muller Algorithm

How to generate Normal Distribution?

$X = N(0,1)$ $\quad Y = N(0,1)$ $\quad P(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

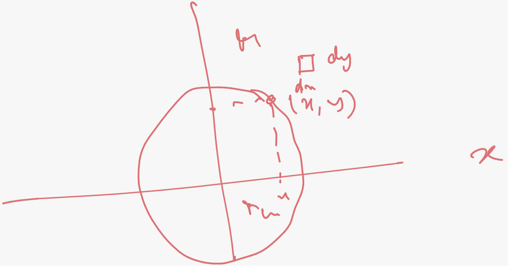
$P(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$

$\boxed{P(x,y) = P(x)\,P(y)}$

$$\hat{P}(x,y) \longrightarrow \int_{-\infty}^{\infty} e^{-x^2/2}\,dx \int_{-\infty}^{\infty} e^{-y^2/2}\,dy = ?$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2/2}\, e^{-y^2/2}\,dx\,dy \longleftarrow$$

$$\hat{p}(x,y) : \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\, dx\, dy$$

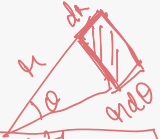


$$r^2 = x^2 + y^2$$

$$\theta$$

$$dx\, dy = r\, dr\, d\theta$$

$$\hat{p}(x,y) : \int_0^{2\pi}\int_0^{\infty} e^{-r^2/2}\, r\, dr\, d\theta$$



$$\boxed{2\pi \int_0^{\infty} e^{-r^2/2}\, r\, dr}$$

$$\hat{P}(x,y) = 2\pi \int_0^\infty e^{-r^2/2} \, r\,dr$$

$$z = \frac{r^2}{2} \qquad dz = r\,dr$$

$$= 2\pi \int_0^\infty e^{-z} \, dz$$

$$= 2\pi \left[-e^{-z} \right]_0^\infty = 2\pi$$

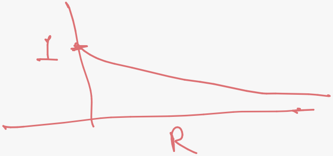
$$\hat{P}(x,y) = 2\pi$$

$X \sim N(0,1)$ $\quad$ $Y \sim N(0,1)$

$\rightarrow U_1 \sim \text{Unif}[0,1]$ $\quad$ $U_2 \sim \text{Unif}[0,1]$

$$P(R,\theta) = \frac{1}{2\pi} e^{-R^2/2} \longleftrightarrow P(x,y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$

$$U_1 = \frac{\Theta}{2\pi} \qquad \underline{U_2 = e^{-R^2/2}}$$



$$X = R\cos\Theta = \sqrt{-2\ln U_2}\,\cos(2\pi U_1)$$

$X \sim N(0,1)$ $\quad$ $Y = R\sin\Theta$

Box-Muller Algorithm

How to generate Normal Distribution?

Reference

Rejection Sampling

Sampling from a complicated distribution or in higher dimensional space

Rejection Sampling

Sampling from a complicated distribution or in higher dimensional space

- Uniform on A where A is complicated!

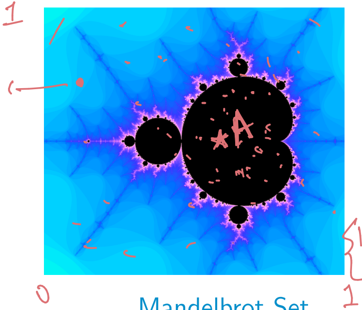
Rejection Sampling

Sampling from a complicated distribution or in higher dimensional space

- Uniform on A where A is complicated!

A C B

$\frac{1}{A}$



Mandelbrot Set

Y

$\min(Y) = \boxed{\frac{A}{1}}$

$Y_1 \quad Y_2 \quad \ldots \quad Y_n$

$\frac{10000}{A}$

Rejection Sampling

Uniform Distribution

- If $A \subset B$

Rejection Sampling

Y_1 Y_2 Y_3 Y_4 . . .

$(0,0)$ $(0,1)$ $(0.1,1)$ . . .

$(0,1)$ $(0,0)$

Uniform Distribution

- If $A \subset B$
- $Y_1, Y_2, \ldots$ are independently and uniformly drawn from B

Rejection Sampling

$\times \quad \times \quad A \quad A \ \ldots$

$Y_1 \quad Y_2 \quad Y_3 \ \ldots \quad Y_n$

Uniform Distribution

- If $A \subset B$
- $Y_1, Y_2, \ldots$ are independently and uniformly drawn from B
- $X = Y_K$ where $K = \min\{k : Y_k \in A\}$

R.V

$\rightarrow X = Y_3$

$\boxed{X \sim \text{Unif}(A)}$

Unf $\{1, 2, 3, 4, 5\}$ $\rightarrow$ Unif $\{2, 3, 4\}$

$Y_1 \quad Y_2 \quad Y_3 \quad Y_4 \ \ldots \ Y_8$

$1 \quad 1 \quad 2 \quad 2 \ \ldots$

$\rightarrow 1 \quad 4 \quad 5 \quad 2 \ \ldots$

$X = 2$

$X = 2$

$X = 3$

Rejection Sampling

Uniform Distribution

- If $A \subset B$
- $Y_1, Y_2, \ldots$ are independently and uniformly drawn from B
- $X = Y_K$ where $K = \min\{k : Y_k \in A\}$
- Then $X \sim Unif(A)$

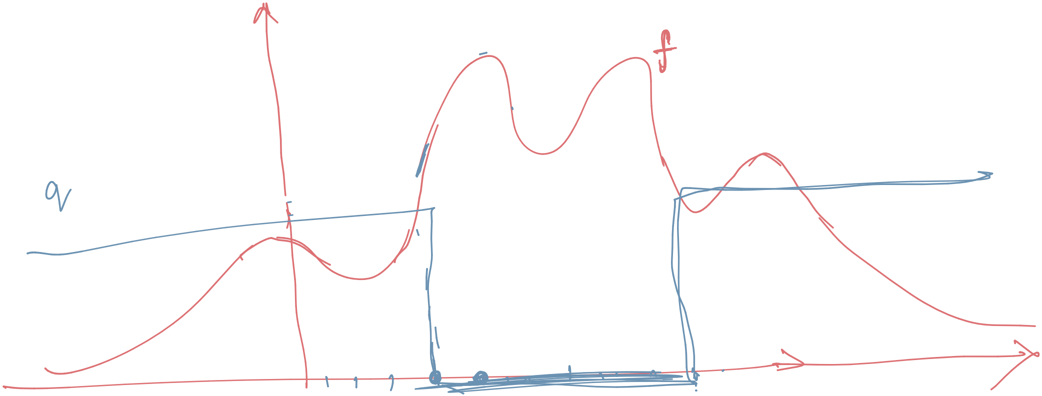
Rejection Sampling

Uniform Distribution

- If $A \subset B$
- $Y_1, Y_2, \ldots$ are independently and uniformly drawn from B
- $X = Y_K$ where $K = \min\{k : Y_k \in A\}$
- Then $X \sim Unif(A)$
- Note that K is a random variable

Rejection Sampling

Sampling from a complicated Non-uniform distribution f on $\mathbb{R}^d$



Rejection Sampling

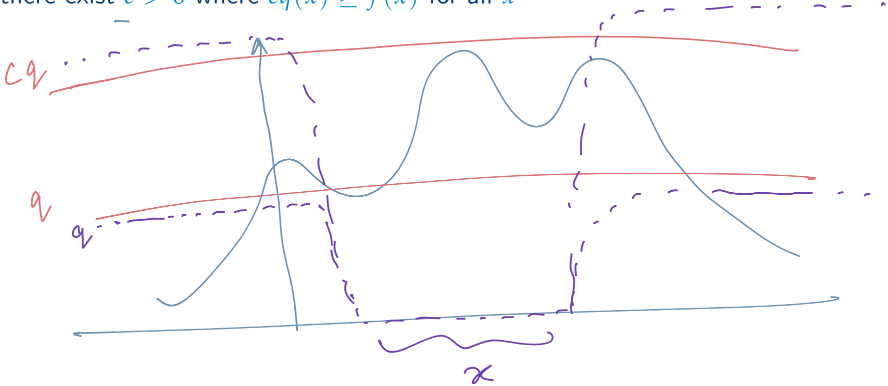
Sampling from a complicated Non-uniform distribution f on $\mathbb{R}^d$

- Choose a proposal distribution q

Rejection Sampling

Sampling from a complicated Non-uniform distribution f on $\mathbb{R}^d$

- Choose a proposal distribution q
  - there exist $c > 0$ where $cq(x) \geq f(x)$ for all x



Rejection Sampling

Sampling from a complicated Non-uniform distribution f on $\mathbb{R}^d$

- Choose a proposal distribution q
  - there exist $c > 0$ where $cq(x) \geq f(x)$ for all x
  - Easy to sample from

Rejection Sampling

Sampling from a complicated Non-uniform distribution f on $\mathbb{R}^d$

- Choose a proposal distribution q
  - there exist $c > 0$ where $cq(x) \geq f(x)$ for all x
  - Easy to sample from
- Sample from $X \sim q$

Rejection Sampling

Sampling from a complicated Non-uniform distribution f on $\mathbb{R}^d$

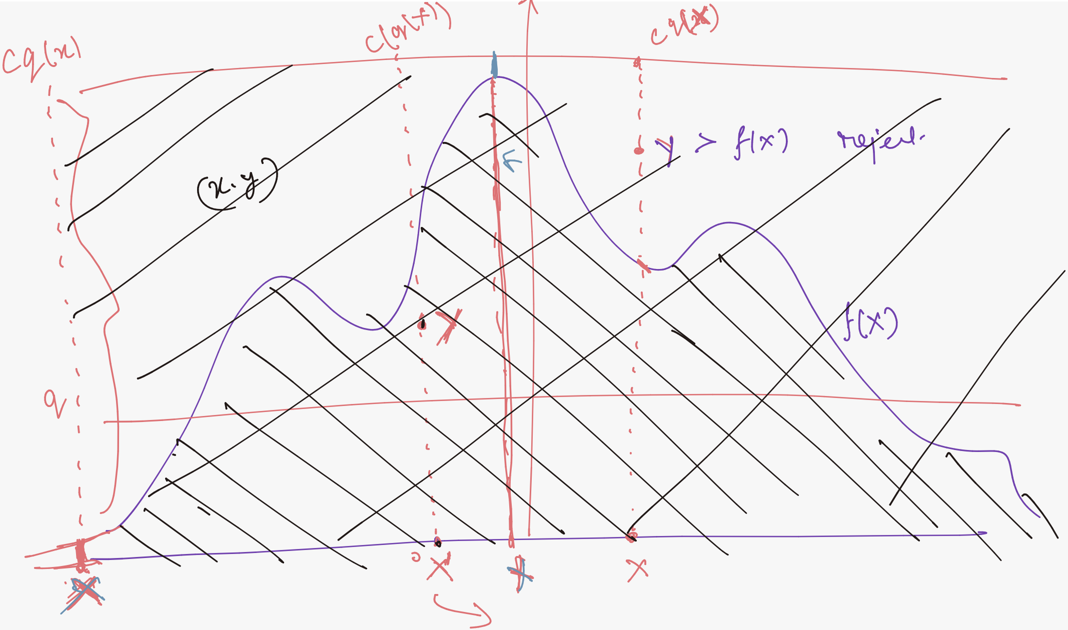
- Choose a proposal distribution q
  - there exist $c > 0$ where $cq(x) \geq f(x)$ for all x
  - Easy to sample from
- Sample from $X \sim q$
- Given X, we sample $Y \sim Unif(0, cq(X))$

Rejection Sampling

Sampling from a complicated Non-uniform distribution f on $\mathbb{R}^d$

- Choose a proposal distribution q
  - there exist $c > 0$ where $cq(x) \geq f(x)$ for all x
  - Easy to sample from
- Sample from $X \sim q$
- Given X, we sample $Y \sim Unif(0, cq(X))$
- Reject as follows

$$\begin{cases} Y \leq f(X), & Z = X \\ \text{reject} \end{cases}$$



cq(x)
c(q(f))
cq(x)
(x, y)
y > f(x) reject
f(x)
q
x
f
x

Rejection Sampling

Sampling from a complicated Non-uniform distribution f on $\mathbb{R}^d$

- Choose a proposal distribution q
  - there exist $c > 0$ where $cq(x) \geq f(x)$ for all x
  - Easy to sample from
- Sample from $X \sim q$
- Given X, we sample $Y \sim Unif(0, cq(X))$
- Reject as follows

$$\begin{cases} Y \leq f(X), & Z = X \\ \text{reject} \end{cases}$$

- $Z \sim f$

Rejection Sampling (Intuition)

- Rejecting uniformly the area about f and below cq

Rejection Sampling (Intuition)

- Rejecting uniformly the area about f and below cq
- Choosing a good q

Rejection Sampling

- If $X \sim q$ and $Y|X \sim Unif(0, cq(x))$ then $(X, Y) \sim Unif(B)$ where $B = \{(x, y) : x \in \mathbb{R}^d,\ 0 < y < cq(x)\}$

Rejection Sampling

- If $X \sim q$ and $Y|X \sim Unif(0, cq(x))$ then $(X, Y) \sim Unif(B)$ where $B = \{(x, y) : x \in \mathbb{R}^d,\ 0 < y < cq(x)\}$
- Viz

Expectation

$$\mathbb{E}[X] = \begin{cases} \sum_x xf(x); & X \text{ is discrete} \\ \int_{-\infty}^{\infty} xf(x)dx; & X \text{ is continuous} \end{cases}$$

Expectation

$$\mathbb{E}[X] = \begin{cases} \sum_x x f(x); & X \text{ is discrete} \\ \int_{-\infty}^{\infty} x f(x) dx; & X \text{ is continuous} \end{cases}$$

Linearity: $\mathbb{E}[aX + Y] = a\mathbb{E}[X] + \mathbb{E}[Y]$

Examples

Find the expectation of

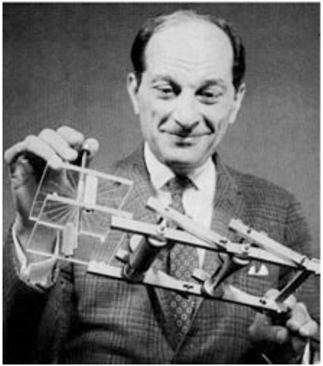
- Y, $f(y) = \frac{y+2}{25}$; $y = \{1, 2, 3, 4, 5\}$
- $X \sim \text{Geometric}(\lambda)$

Monte Carlo Sampling

How to approximate expectations?

- without access to the distribution's closed form
- sum or integral is intractable

Monte Carlo Sampling



Monte Carlo Sampling

Goal: Approximate $\mathbb{E}[f(X)]$

Monte Carlo Sampling

Goal: Approximate $\mathbb{E}[f(X)]$

- If $X_1, X_2, \ldots, X_n$ are sampled from a distribution p

Monte Carlo Sampling

Goal: Approximate $\mathbb{E}[f(X)]$

- If $X_1, X_2, \ldots, X_n$ are sampled from a distribution p
- $X_i's$ are independent and identically distributed

Monte Carlo Sampling

Goal: Approximate $\mathbb{E}[f(X)]$

- If $X_1, X_2, \ldots, X_n$ are sampled from a distribution p
- $X_i's$ are independent and identically distributed
- The approximate expectation is given by,

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

Monte Carlo Sampling

Goal: Approximate $\mathbb{E}[f(X)]$

- If $X_1, X_2, \ldots, X_n$ are sampled from a distribution p
- $X_i's$ are independent and identically distributed
- The approximate expectation is given by,

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

- $\hat{\mu}$ is an estimator for $\mathbf{E}[f(X)]$

Unbiased Estimator

The sample mean $\hat{\mu}$ is an

- Unbiased estimator

$$\mathbb{E}[\hat{\mu}] = \mathbb{E}[f(X)]$$

Unbiased Estimator

The sample mean $\hat{\mu}$ is an

- Unbiased estimator

$$\mathbb{E}[\hat{\mu}] = \mathbb{E}[f(X)]$$

- Consistent estimator ($\sigma^2(f(X)) < \infty$)

$$\Pr(|\hat{\mu}_n - \mathbb{E}[f(X)]| < \epsilon) \to 1$$

Unbiased Estimator

The sample mean $\hat{\mu}$ is an

- Unbiased estimator

$$\mathbb{E}[\hat{\mu}] = \mathbb{E}[f(X)]$$

- Consistent estimator ($\sigma^2(f(X)) < \infty$)

$$\Pr(|\hat{\mu}_n - \mathbb{E}[f(X)]| < \epsilon) \to 1$$

- $\frac{1}{n}\sigma^2(f(X)) \to 0$