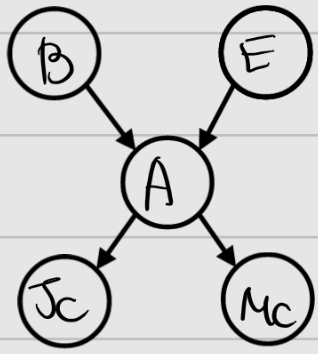
< Forward Sampling >

$P(B)=0.001$



$P(E)=0.002$

B	E	P(A|B,E)
T	T	0.95
T	F	0.94
F	T	0.29
F	F	0.001

A	P(Jc|A)
T	0.9
F	0.05

A	P(Mc|A)
T	0.7
F	0.01

topological order 따름

① B=False

② E=False

③ $A \mid B=F, E=F$ → True

④ Jc|A → True

⑤ Mc|A → True

→ ①~⑤과 같은 순서로 수많은 sample 생성

$P(E=T \mid MC=T) = ?$ → Counting하여 계산하기

문제점: 오차 존재 (n→∞ 오차무시 가능?)

↳ 시간

<Rejection Sampling>

Q. $P(E=T \mid Mc=T, A=F) = ?$

상황 ① B=False ok

② E=True ok

③ A|B,E = False → 오류 문제조건에 A=F 이므로

↳ given 조건에 맞지 않아서 reject 이후 처음부터 sampling

∴ 상황과 given 비교하며 reject 여부 결정

$$f(x) = 0.3e^{-0.2x^2} + 0.7e^{-0.2(x-10)^2}$$

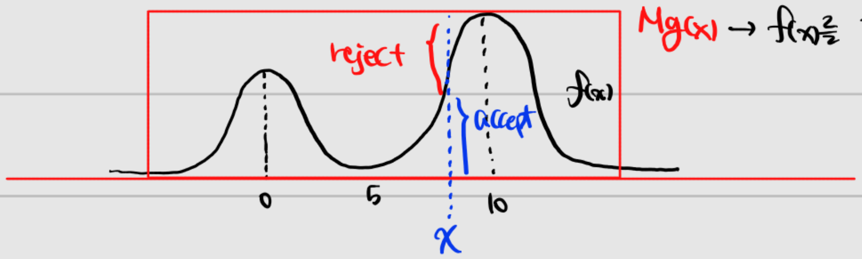
g(x) : 제안분포 → 쉽게 sample 추출

ex) normal distribution, uniform distribution ...

g(x): Uniform distribution

$$g(x) = \begin{cases} \frac{1}{24} & -7 \le x < 17 \\ 0 & \text{otherwise} \end{cases}$$

⇓



Mg(x) → f(x)를 감쌀 수 있어야 함

$\frac{f(x)}{Mg(x)}$ > random → accept

< → reject

<Importance Sampling>

단점: rejection → waste ↑

$$E_{x\sim p}[f(x)] = \int f(x)\,p(x)\,dx = \int \boxed{f(x) \cdot \frac{p(x)}{q(x)}} \cdot q(x)\,dx$$

$$= E_{x\sim q}\left[f(x) \cdot \frac{p(x)}{q(x)}\right]$$

$$= \frac{1}{N} \sum_{n=1}^{N} f(x_n) \cdot \frac{p(x_n)}{q(x_n)}$$

$\frac{p(x_n)}{q(x_n)} \rightarrow r_n$ (importance weight)

위는 continuous

discrete domain 의 경우는....?

$P(E \mid Mc=T, A=F) = ?$

Sum sw = Normal sw = 0

Sampling weight = 1 (sw)

① B = False

② E = False

③ $A=F \mid B=F, E=F$

$P(A=F \mid B=F, E=F) = 0.999$

$sw = 1 \times 0.999 = 0.999$

④ $Jc=T \mid A=F$

⑤ $Mc=T \mid A=F$

$P(Mc=T \mid A=F) = 0.01$

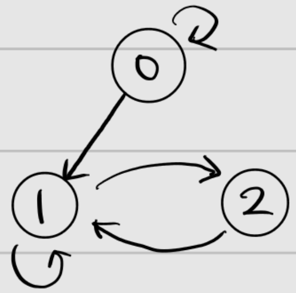
$sw = 0.999 \times 0.01$ → If $E=T$

Sum sw += sw

normal sw += sw

return Sum sw / normal sw

Markov Chain 특성



Reducible

1,2 → 0 불가능

⓪ : Transcience

0→1로 가면 다시 0으로 올 수 없음

① ② : recurrent → 다시 돌아오기 가능

irreducible → $\forall i \in S,\ \forall j \in S \quad i \leftrightarrow j$

모든 i,j가 서로 communicate 가능

- Periodicity

$$d = \gcd\{n : T_{ij}^{n} > 0\}$$

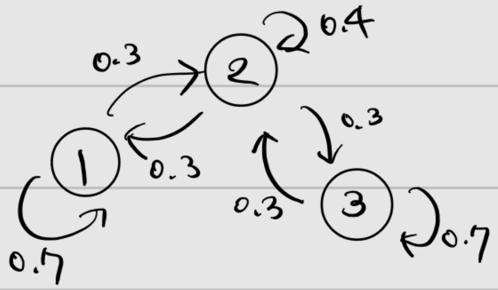
State (i)에 2회, 3회, 4회, 7회에 recurrent

→ gcd=1

if d=1 → aperiodic

- Ergodicity : recurrent + aperiodic

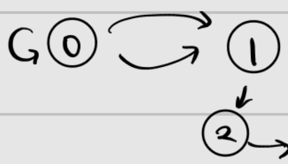
<Stationary distribution>



$$T = \begin{pmatrix} 0.7 & 0.3 & 0 \\ 0.3 & 0.4 & 0.3 \\ 0 & 0.3 & 0.7 \end{pmatrix}$$

$$\pi = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

$\pi T = \pi$ 를 만족



→ absorbing state (순환X)

Markov Chain irreducible & ergodic

$RT_i = \min\{n>0\ x_n = i \mid x_0 = i\}$

return time to state i

$$\pi_i = \lim_{n\to\infty} T_{ij}^{(n)} = \frac{1}{E[RT_i]}$$

→ π_i는 Unique 하게 결정된다

T 계산

$$\pi\,(\underset{(s\times s)}{I} - T + \underset{(s\times s)}{I}) = 1_{(1\times s)}$$

$\pi - \pi T + \pi = 1$ 꼴

↳ $\pi_j = \sum_{i\in S} \pi_i T_i$ $\quad$ $\sum_{i\in S} \pi_i = 1$

Reversible MC vs Irreversible MC

reversible : $\pi_i T_{ij} = \pi_j T_{ji}$

즉, $i\to j$ & $j\to i$ 확률이 동일

Reversible Markov Chain $\xrightarrow[\text{만족}]{}$ Stationary distribution

역은 성립 X

< MCMC (Markov Chain Monte Carlo) >

Monte Carlo : 무한한 시도 → 정답   현실? 불가능...

→ 유한한 시도로 정답을 추정

Current value: z^t

Propose candidate $z^* \sim q(z^*|z^t)$ q^t: proposal distribution

acceptance probability α

$$\begin{cases} \text{Accept} : z^{t+1} = z^* \\ \text{Reject} : z^{t+1} = z^t \end{cases}$$

(Metropolis-Hastings Algorithm)

$P(\theta) \propto g(\theta)$ 정규화 때문에 해줌

① Select θ_0

② for $i = 1, 2 \cdots m$:

a) 그리기 $\theta^* \sim q(\theta^*|\theta_{i-1})$ q: proposal distribution

b) $$\alpha = \frac{g(\theta^*) / q(\theta^*|\theta_{i-1})}{g(\theta_{i-1}) / q(\theta_{i-1}|\theta^*)} = \frac{g(\theta^*) \cdot q(\theta_{i-1}|\theta^*)}{g(\theta_{i-1}) \cdot q(\theta^*|\theta_{i-1})}$$

c) $\alpha \geq 1$ accept θ^*, set $\theta_i = \theta^*$

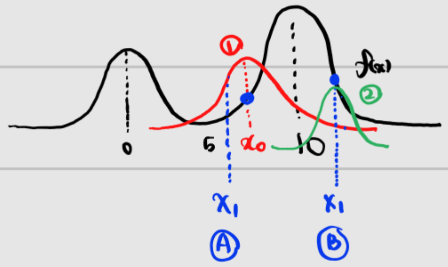
$0 < \alpha < 1$ accept θ^*, set $\theta_i = \theta^*$ with prob α

else   reject θ^*, set $\theta_i = \theta_{i-1}$ with prob $(1-\alpha)$

① Random Initialize → 하나의 random 선택 $x_0 = 7$

$f(x) = 0.3e^{-0.2x^2} + 0.7e^{-0.2(x-10)^2}$

② x_0를 중심으로 제안 분포 → 현재는 정규분포 선택



A,B → random again

A의 경우 → reject ——→ 기회를 한번 더 준다.

B의 경우 → accept   if) $0 < \frac{f(x_1)}{f(x_0)} < 1$ → 기회 O

$\frac{f(x_1)}{f(x_0)} \geq 1$ → accept

왜 reject 해도 기회를 주는가?

→ 한쪽으로 accept 하면 그 방향으로 몰림 → 편향

<Gibbs Sampling>

여러개의 z (latent) 중 하나씩 update

$p(x_1, x_2, x_3)$

1) 임의의 표본 $X_0 = (x_1^0, x_2^0, x_3^0)$

2) 변수 하나씩만 변경

a) X_0의 x_2^0, x_3^0는 고정 → x_1^0를 update

$p(x_1^1 | x_2^0, x_3^0)$

b) x_1^1, x_3^0 고정 → x_2^1 update

c) x_1^1, x_2^1 → x_3^1 update

$X_1 = (x_1^1, x_2^1, x_3^1)$

$X = \{1, 2, 3\}$ $\qquad$ $Y = \{4, 5, 6\}$

$P(X=1) = \frac{1}{6}$ $\qquad$ $P(Y=4) = \frac{1}{6}$

$P(X=2) = \frac{1}{3}$ $\qquad$ $P(Y=5) = \frac{1}{3}$

$P(X=3) = \frac{1}{2}$ $\qquad$ $P(Y=6) = \frac{1}{2}$

X, Y는 종속)

	X=1	X=2	X=3
Y=4			0.1
Y=5		0.1	0.2
Y=6	0.1	0.2	0.3

X=1	X=2	X=3
$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$

이미 차이O

(종속이므로)

즉, X 먼저 고려 → Y 고려하면

X=1, X=2가 뽑힐 확률이 전혀 달라짐

따라서

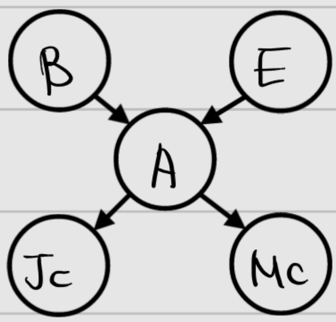
$X_0 \sim P(X)$

$Y_0 \sim P(X \mid X = X_0)$

$X_1 \sim P(X \mid Y = Y_0)$

$Y_1 \sim P(X \mid X = X_1)$

$\vdots$



만약 A=F, Mc=T given

Markov Blanket 내부만 고려

ex) B=? ⇒ E, A만 고려