· 몬테카를로 방법 (Monte Carlo Method)

: 몬테카를로 방법(Monte Carlo Method)이란, 랜덤 표본을 뽑아 함수의 값을 확률적으로 계산하는 알고리즘
수학이나 물리학에서 자주 사용되며 계산하려는 값이 닫힌 형식으로 표현되지 않거나, 복잡한 경우에
그 값을 근사적으로 계산할때 사용됨.

ex) 4분원 원주율 계산

: 반지름이 1 인 4분원의 원주율 구하기

① [0,1], [0,1] 에서 점 (x, y)를 무작위로 뽑는다.

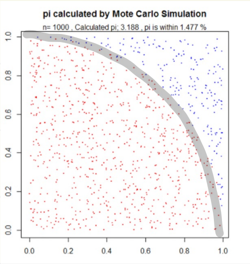
② 이 점이 중심이 (0,0) 이고 반지름이 1인 원에 속하는 지 계산한다.

→ $x^2+y^2 \le 1$ 만족하면 빨간색, 아니면 파란색으로 분류

③ 위 과정 n번 반복

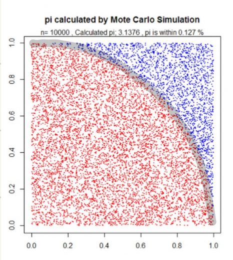
④ 전체 점의 개수를 빨간색 점의 개수로 나눈 비율을 통해 $\frac{\pi}{4}$를 구한다.

(x4 해서 π값 추정하기)



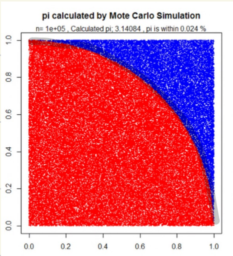
시행횟수 = 1,000 번
계산된 π값 = 3.188

→



시행횟수 : 10,000번
계산된 π값 : 3.13

→



시행횟수 100,000번
계산된 π값 : 3.14

∴ 시행횟수 증가할수록, 계산된 π값은 실제 π값(3.141592 …) 에 근사한다.

• 마코프 연쇄 (Markov chain)

: 마코프 연쇄(Markov chain)란, 마코프 가정(Markov assumption)을 따르는 이산 시간 확률 과정을 가르킨다. 특정 시점의 상태확률은 단지 그 이전 상태에만 의존한다는 것이 핵심.

즉, 한 상태에서 다른 상태로의 전이(transition)는 그동안 상태 전이에 대한 긴 이력(history)을 필요로 하지 않고, 바로 직전 상태에서의 전이로 추정할 수 있다.

$$\Rightarrow P(q_i \mid q_1, \cdots, q_{i-1}) = P(q_i \mid q_{i-1})$$

그런데, 특정 조건을 만족한 상태에서 마코브 연쇄를 반복하다보면 현재상태의 확률이 직전상태와 같아지게 된다. (수렴) 이렇게 평형 상태에 도달한 확률 분포를 정적 분포(Stationary Distribution) 라고 한다.

cf. 전이확률 : 현재 시점(단계)의 상태에서 다음시점의 상태로 될 확률

전이 행렬 : 특정 상태들이 가질 수 있는 모든 전이 확률을 행렬로 만든 것.

상태확률 : 특정 시점에서 발생하는 상태의 확률. 상태확률의 합 = 1

특정시점의 상태확률 = (시작 시점의 상태확률) X (특정 시점의 전이행렬)

안정 상태 : 상태확률에 전이확률을 곱해도 상태 확률이 변하지 않는 상태

ex) $A_A(i)$: 현재 A맥주 선호하는 소비자 i번째 기간에도 A맥주 선호확률 / $B_A(i)$: 현재 A맥주 선호, i번째 기간에는 B 맥주 선호 확률

$A_B(i)$: 현재 B맥주 선호. i기간에는 A맥주 선호 확률 / $B_B(i)$: 현재 B맥주 선호, i기간에도 B 맥주 선호 확률

⇒ 현재 A맥주 선호 소비자 : $[A_A(1)=1 \quad B_A(1)=0]$, 현재 B맥주 선호 : $[A_B(1)=0 \quad B_B(1)=1]$ 로 정의

현재 A맥주 선호 소비자가 다음달에 선호할 맥주 : $[A_A(2) \quad B_A(2)] = [A_A(1) \quad B_A(1)] \underbrace{\boxed{\begin{bmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix}}}_{\text{전이 행렬}} = [1 \quad 0]\begin{bmatrix} .6 & .4 \\ .2 & .8 \end{bmatrix} = [0.6 \quad 0.4]$ 일 때,

$[A_A(3) \quad B_A(3)] = [A_A(2) \quad B_A(2)]\begin{bmatrix} .6 & .4 \\ .2 & .8 \end{bmatrix} = [0.44 \quad 0.56]$

$[A_A(4) \quad B_A(4)] = [A_A(3) \quad B_A(3)]$ 〃 $= [0.38 \quad 0.62]$

⋮

$[A_A(7) \quad B_A(7)] = [0.34 \quad 0.66]$

$[A_A(8) \quad B_A(8)] = [0.33 \quad 0.67]$ 〉수렴됨

$[A_A(9) \quad B_A(9)] = \boxed{[0.33 \quad 0.67]}$

안정상태하의 확률 (Steady-State probability)

∴ 일정기간 지나면 $[A_A(n) \quad B_A(n)] = [A_A(n+1) \quad B_A(n+1)]$ 성립

$A_A + B_A = 1$ 인 점과 $[A_A \quad B_A] = [A_A \quad B_A]\begin{bmatrix} .6 & .4 \\ .2 & .8 \end{bmatrix}$ 이용해

안정하상태의 확률 쉽게 구할 수 있음.

• 마코프 연쇄 몬테카를로 방법 (MCMC)

: MCMC란, 마코프 연쇄에 기반한 확률분포로부터 표본을 추출하는 몬테카를로 방법. MCMC 알고리즘은 샘플을 얻으려고 하는 목표 분포를 Stationary Distribution으로 가지는 마코프체인을 만든다. 이 체인의 시뮬레이션을 가동하고 초기값에 영향을 받는 burn-in period를 지나고나면 목표 분포를 따르는 샘플이 만들어진다.

• MCMC : 깁스 샘플링

: 깁스 샘플링은 MCMC 일종. 이때, MCMC 와 몬테카를로 방식은 다르다. 몬테카를로 방법은 모든 샘플이 독립 (independent)이고 생성될 (추출될) 확률 또한 랜덤이다. 반면, 마코프 연쇄에 기반한 MCMC는 다음번 추출될 샘플은 현재 샘플의 영향을 받는다. 깁스 샘플링은 다음번에 생성될 표본은 현재 샘플에 영향을 받는다는 점에서 MCMC와 같지만, 나머지 변수는 그대로 두고, 한 변수에만 변화 준다는 점이 다르다.

ex) 3개의 확률 변수 결합 분포 $P(X_1, X_2, X_3)$로부터 1개의 표본을 얻으려고 할 때, 깁스 샘플링 절차:

① 임의의 표본 $X^0 = (x_1^0, x_2^0, x_3^0)$을 선택한다.

② 모든 변수에 대해 변수 하나만을 변경하여 새로운 표본 X^1 뽑는다.

→ 실제 사용시, 처음 수집한 표본 X^0은 버리고 X^1만 쓰게된다.

a. 현재 주어진 표본 X^0의 2nd, 3rd 변수 (x_2^0, x_3^0) 고정.

b. 첫번째 기존 변수 x_1^0를 대체할 새로운 값 x_1^1을 $P(x_1^1 | x_2^0, x_3^0)$ 확률로 뽑는다.

c. 1st, 3rd 변수 고정 (x_1^1, x_3^0) 후 2nd 변수 (x_2^0) 대체할 새로운값 (x_2^1) $P(x_2^1 | x_1^1, x_3^0)$ 확률로 추출

d. 1st, 2nd 변수 (x_1^1, x_2^1) 고정, 3rd 변수 (x_3^0) 대체할 새로운 값 x_3^1을 $P(x_3^1 | x_1^1, x_2^1)$ 확률로 추출

e. 최종적으로 구한 $X^1 = (x_1^1, x_2^1, x_3^1)$

⇒ 새로운 값 추출시 쓰인 조건부확률은 결합확률 분포 $P(X_1, X_2, X_3)$에 비례

표본의 앞부분은 초기 상태 X^0에 크게 의존.

But, 충분히 많이 뽑고 난 뒤에는 초기 상태에 관계없이 P에 기반한 표본 수집 가능