Discrete Probability

확률분포란 (probability distribution) ?

: 확률변수가 특정한 값을 가질 확률을 나타내는 함수를 의미한다.

확률분포는 확률변수가 어떤 종류의 값을 가지는가에 따라 크게 이산확률분포 (discrete probability distribution) 와 연속확률 분포(continuous probability distribution) 중 하나에 속하며, 둘 중 어디에도 속하지 않는 경우도 존재한다.

이산확률분포란 (discrete probability distribution) ?

: 이산확률변수가 가지는 확률분포를 의미한다. 여기에서 확률변수가 이산 확률변수라는 말은 확률변수가 가질 수 있는 값의 개수가 셀 수 있다는 의미이다.

Stochastic process

<u>확률과정이란 (stochastic process) ?</u>

: 확률법칙에 의해 생성되는 일련의 통계적인 현상을 각 시점에서 나타나는 수치적인 양을 확률변수로 기술하고 이 확률변수를 시간 흐름별로 나열한 집합으로 모형화 할 때 사용된다. 여기서 시간의 흐름이 이산적일 때 이산 시간 확률과정 또는 시계열(time series)이라고 하고 연속적일 때는 연속시간 확률과정이라고 한다.

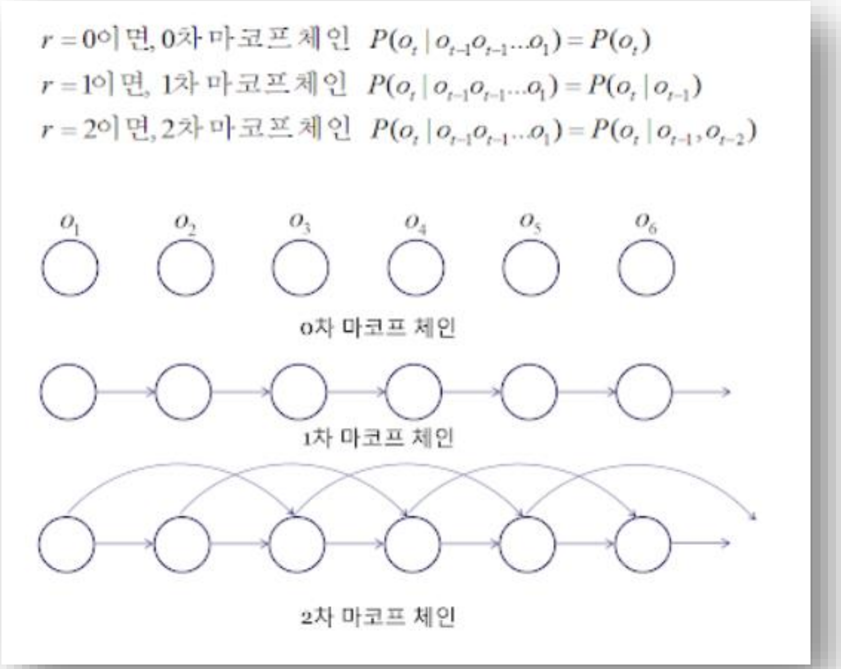
Markov Chain

마르코프 체인은 마르코프 성질(Markov property)을 지닌 이산확률 과정(Discrete Stochastic Process)을 의미함.

[성질] :

시간 (t+1)에서의 상태는 시간 t 혹은 그 이전 일정 기간의 상태에만 영향을 받는다.

즉, 특정 상태의 확률은 오직 과거 상태에 의존한다.



1차 마르코프 체인은 n번째 상태가
n+1번째 상태를 결정하는데에 영향을 미침.

(시간 t에서의 관측은 단지 최근 r개의 관측에만 의존한다는 가정을 하고 그 가정하에 성립)

Markov Model

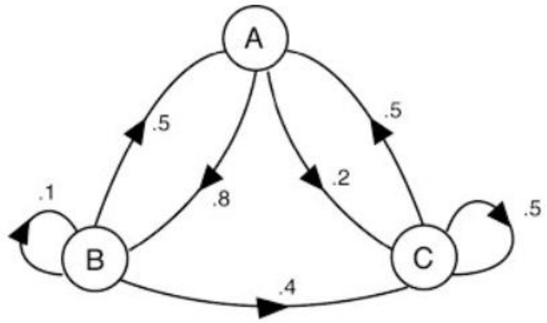
가정 (시간 t에서의 관측은 단지 최근 r개의 관측에만 의존한다는 가정을 하고 그 가정하에 성립) 하에 확률적 모델을 만든 것.

상태 (state) : $V = \{v_1, \dots, v_m\}$ (m개의 상태가 존재)

상태 전이 확률 : a_{ij} (상태 v_i 에서 v_j로 이동할 확률)

$$a_{ij} = P(o_t = v_j | o_{t-1} = v_i)$$

$$a_{ij} > 0 \; and \; \sum_{j=1}^{m} a_{ij} = 1$$



Markov graph of transiton probabilites between states A, B and C

Markov Model

오늘/내일	맑음	비	흐림
맑음	0.8	0.05	0.15
비	0.2	0.6	0.2
흐림	0.2	0.3	0.5

오늘 날씨(q1)가 “맑음”일 경우, 내일 날씨(q2)가 “맑음 " 이 되고 모레 날씨(q3)가 “비 " 가 될 확률

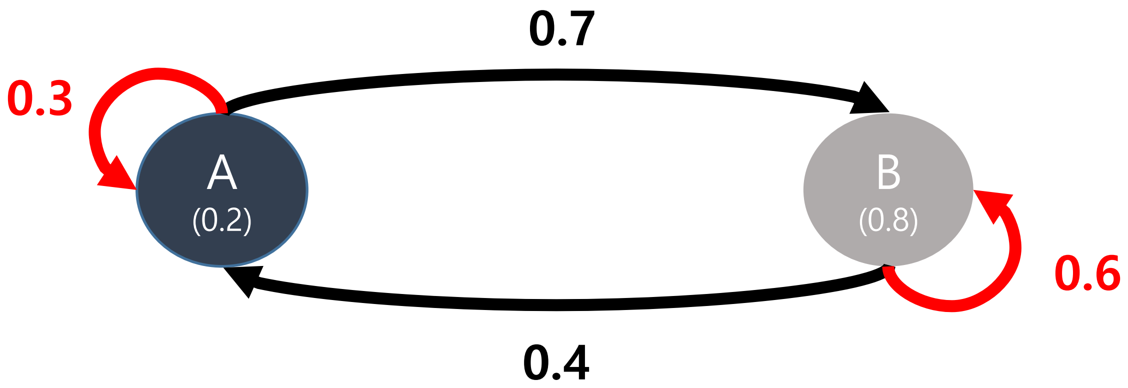
$$P(q2 = \mathrm{sunny}, q3 = rainy \mid q1 = sunny)$$

$$= P(q3 = rainy \mid q2 = sunny, q1 = sunny) * P(q2 = sunny \mid q1 = sunny)$$

$$= P(q3 = rainy \mid q2 = sunny) * P(q2 = sunny \mid q1 = sunny)$$

$$= 0.05 * 0.8$$

$$= 0.04$$



$$X(0) = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}$$

Transition matrix (P)

$$X(1) = \begin{pmatrix} (0.3)(0.2) + (0.4)(0.8) \\ (0.7)(0.2) + (0.6)(0.8) \end{pmatrix} = \begin{pmatrix} 0.3 & 0.4 \\ 0.7 & 0.6 \end{pmatrix} \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix} = P * X(0)$$

Transition matrix (P)

$$X(2) = \begin{pmatrix} (0.3)(0.38) + (0.4)(0.62) \\ (0.7)(0.38) + (0.6)(0.62) \end{pmatrix} = \begin{pmatrix} 0.3 & 0.4 \\ 0.7 & 0.6 \end{pmatrix} \begin{pmatrix} 0.38 \\ 0.62 \end{pmatrix} = P * X(1) = P^2 * X(0)$$

$P(t+1$에 B노드 | t에 A노드) ← 조건부 확률

t 시점에서 A일 확률이었다가 t+1 시점에서 B가 될 확률

$$X(t+1) = P * X(t)$$

$$\mathrm{X(t)} = P^t * X(0)$$

Stationary distribution

State matrix X(0), X(1), X(2),,, 은 어떤 행렬 X로 수렴한다.

이 행렬 X는 transition matrix (P)를 곱해도 변하지 않는다. ➔ PX = X

이때 행렬 X는 마르코프 체인의 stationary distribution 을 나타낸다.

[구하는 방법]

1. X(0), X(1), X(2),,, 를 계속 구해본다. (Transition matrix (P)를 계속 곱해본다)
2. $\mathrm{X(t)} = P^t * X(0)$ 임을 이용한다. 대각화로 P^t 를 계산한 후 $X(t)$ 를 구한다.
3. $PX = X$ 임을 이용한다. ➔ $(P - I)X = 0$