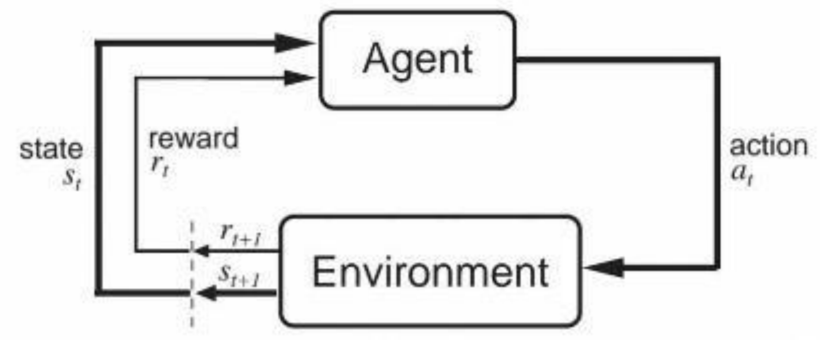
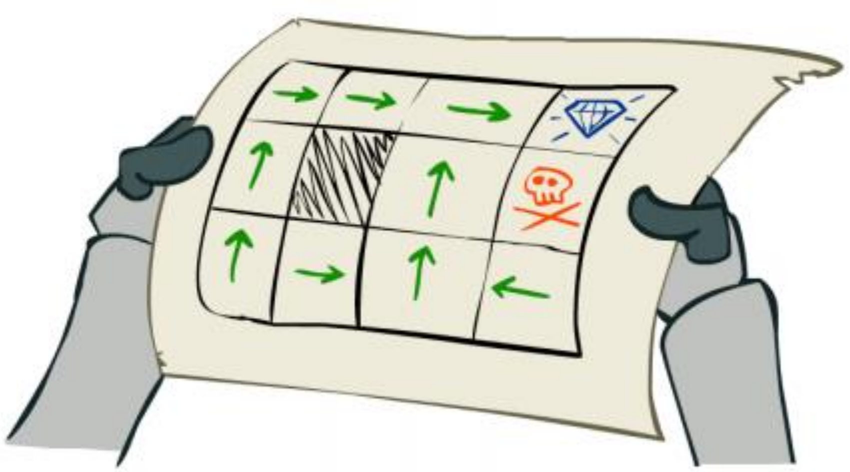
마르코브 프로세스 개요

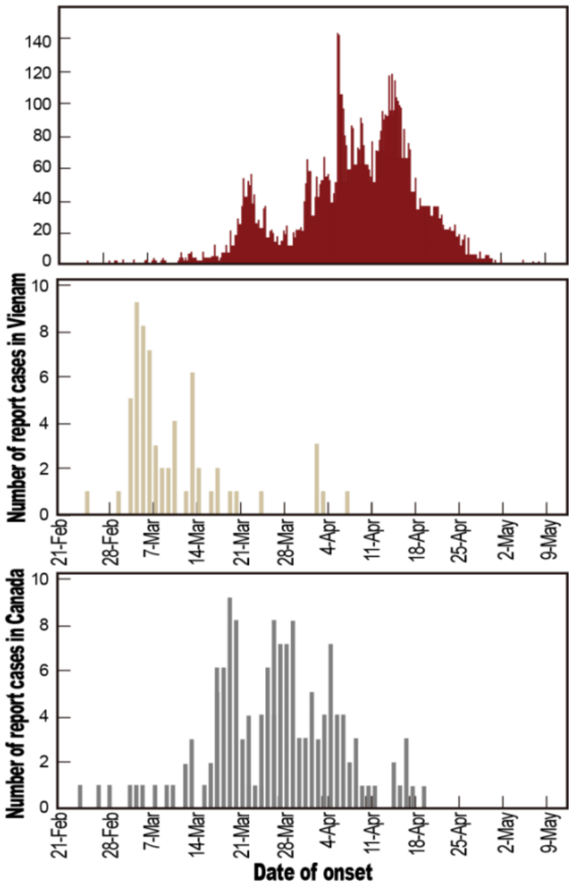




불확실성 모델링

- **확률**
  - 주위에서 발생하고 있는 여러 사건들은 근본적으로 불확실성을 내포
  - 확률은 불확실성을 표현할 수 있는 수단
  - 확률변수와 확률분포
    - 예: 날씨
      - X = 1 (맑음), 2 (흐림), 3 (비)
      - P(X=1) = 0.6, P(X=2) = 0.2, P(X=3)=0.2
  - 시간에 따른 날씨의 시계열
    - 맑음 → 흐림 → 흐림 → 비 → 비 → 맑음 → …
    - 맑음 → 맑음 → 흐림 → 비 → 맑음 → ….

도입예시



Source: https://www.kent.ac.uk/smsas/personal/lb209/files/notes1.pdf

- **전염병 (SARS, CORONA 등)**
  - 주차 별 신규확진자수, 확진율
- **이를 통해 알고 싶은 것은 어떤 것들이 있을까?**
  - 현재까지의 확진자 수 이력을 바탕으로 다음 주에 추가 신규 확진자가 50건 이상 될 확률은?
  - 평균적으로 몇 주 후에 확산세가 진정될 것으로 예상되는가?

- 시시각각 변하는 주식가격, 특정 지점에서의 바람의 세기, 센서의 노이즈 등 시간에 따라 확률적으로 변화하는 프로세스는 다양하다.

확률과정 (추계적 과정)

- **확률과정 (추계적 과정; Stochastic process)**
  - 불확실성을 가지고 변하는 일련의 과정
  - 시간에 따라 어떤 사건이 발생할 확률값이 변화하는 과정
  - 시간에 따른 확률변수들의 집합 $\{X_t: t \in T\}$
    - 시간공간 (time space) T: 관찰시점들의 집합
    - 결합확률분포 $P(X_0), P(X_0, X_1), P(X_0, X_1, X_2), \ldots$
- **상태공간 (State space)**
  - 상태 (State): 어떤 시점에서 확률과정이 갖는 값
    - 신규확진자수, 주가, 바람세기 등
  - 상태공간: 확률과정 $\{X_t: t \in T\}$의 확률변수 X_t 가 가질 수 있는 모든 가능한 값들의 집합
- **결국 확률과정이란 시간이 진행 함에 따라 상태가 확률적으로 변화하는 과정을 의미**

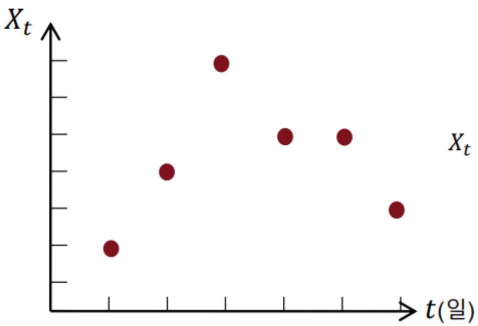
확률과정: 시간공간 T와 상태공간 S에서 정의된 확률변수들의 집합

확률과정 (추계적 과정) 예시

- **(예시) 동전 던지기 게임**
  - 게임 개요
    - 동전 하나를 던져서 앞면이 나오면 +1점, 뒷면 나오면 -1점을 얻어 기존 점수에 더한 것이 동전던지기 게임의 점수
    - 플레이어가 종료를 선언한 시점에서 게임 종료
    - 마지막에 점수가 양수면 점수 당 천원을 받고 음수는 점수 당 천원을 내는 게임
    - 점수는 0부터 시작

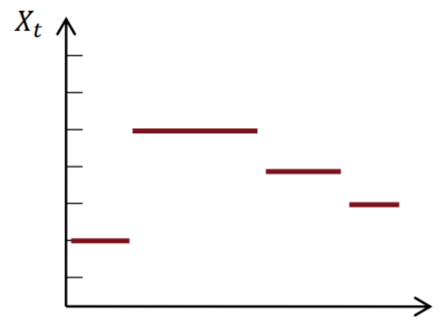
  - 확률과정
    - 확률변수 X_n: n번째 동전을 던졌을 때의 점수
    - 상태공간 $S = \{\dots, -2, -1, 0, 1, 2, \dots\}$
    - 시간공간 $T = \{0,1,2, \dots\}$

확률과정 (추계적 과정) 분류



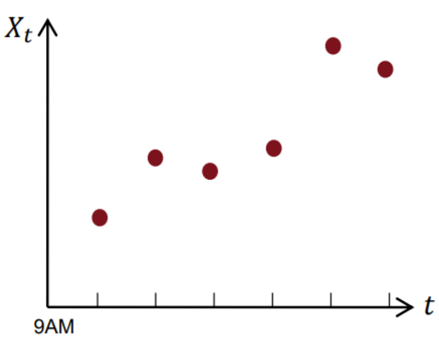
$X_t = t$ 일의 OO회사 주식 종가

이산시간-이산상태 확률과정



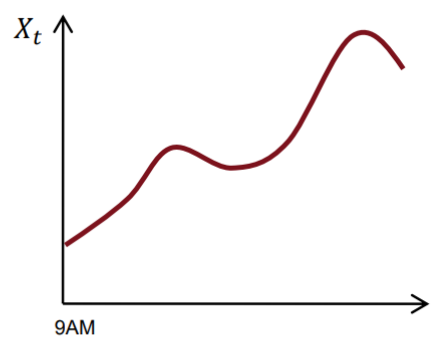
X_t = t시점에서의 OO회사 주가

연속시간-이산상태 확률과정



X_t = 매 t시의 실내 온도

이산시간-연속상태 확률과정



X_t = t시점의 실내 온도

연속시간-연속상태 확률과정

확률과정 (추계적 과정)

- **샘플경로, 에피소드(Sample path; episode)**
  - 시간 t가 변함에 따라 확률과정 $\{X_t : t \in T\}$가 갖는 값의 자취
  - 확률과정의 실현치들을 모아 놓은 것 (일종의 시계열 데이터 샘플)
  - 확률과정이 반복될 때 마다 다른 샘플경로를 보임
    - 동전던지기
      - 0→+1→0→+1→+2→…
      - 0→+1→+2→+1→0→-1→…

- **정상과정 (Stationary process)**
  - 시간 t에 따라 확률법칙이 변하지 않는 확률과정
  - t가 변하더라도 상태확률이 변하지 않는 확률과정

마르코브 성질

- **마르코브 성질 (Markov property)**
  - “미래는 현재로부터 정해지며, 과거는 영향을 주지 못한다”

P(the future | the present, the past)
= P(the future | the present)

$$P(X_{k+1}|X_k, X_{k-1}, \ldots, X_0) =$$

Andrey Markov

- **마르코브 프로세스 (Markov process)**
  - 마르코브 성질을 가진 확률과정

이산시간 마르코브 체인

- **이산시간 마르코브 체인 (Discrete-time Markov chain; DTMC)**
  - 다음의 성질을 만족하는 확률과정 $\{X_n : n \geq 1\}$
    - 모든 n에 대해 $X_n \in S$
    - 모든 $i, j \in S$에 대해,

$$P(\underbrace{X_{n+1} = j}_{\text{미래}} | \underbrace{X_n = i}_{\text{현재}}, \underbrace{X_{n-1}, \ldots, X_0}_{\text{과거}}) = \boxed{P(X_{n+1} = j | X_n = i)}$$

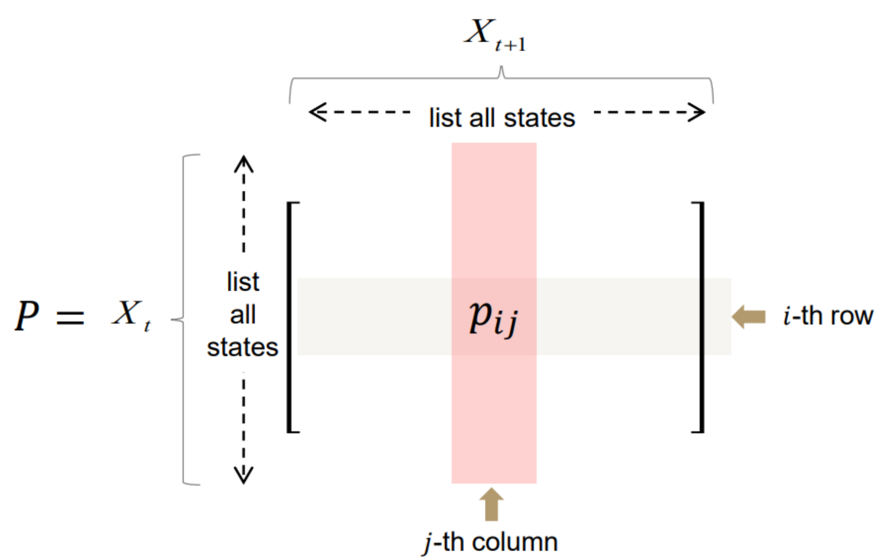
상태 전이확률
(1-step transition probability)

이산시간 마르코브 체인

- **시간 동질 (time-homogeneous) DTMC**
  - 전이확률 $P(X_{n+1} = j | X_n = i)$이 n에 독립적임
    - $P(X_1 = j | X_0 = i) = P(X_{400} = j | X_{399} = i) = P(X_{n+1} = j | X_n = i)$ 성립

$$\Rightarrow \; P(X_{n+1} = j | X_n = i) = p_{ij}$$

(1-단계) 전이 행렬:



이산시간 마르코브 체인

- **시간 동질 (time-homogeneous) DTMC**
  - 전이확률 $P(X_{n+1}=j|X_n=i)$이 n에 독립적임
    - $P(X_1=j|X_0=i)=P(X_{400}=j|X_{399}=i)=P(X_{n+1}=j|X_n=i)$ 성립

$$\Rightarrow\; P(X_{n+1}=j|X_n=i)=p_{ij}$$

(1-단계) 전이 행렬: