Lecture 1 : Introduction

Part 1°. Stochastic Process :

$(X_t)_{t\in T}$	Collection of Random Variables indexed by T
T	index set usually $T \subseteq \mathbb{N} = \{0, 1, 2, \cdots\}$. i.e., " discrete in time "
$X_t : \Omega \longrightarrow \mathcal{X}$	Random Variables (Formally, $\exists (\Omega, \mathcal{F}, P)$. s.t. $\mathbb{P}(X\in S) = \mathbb{P}(\{\omega\in\Omega \mid X(\omega)\in S\})$)
$\mathcal{X}$	State space usually, $\mathcal{X} \subseteq \mathbb{R}$.

Ex 1. $X_1, X_2, \cdots, X_n$ are i.i.d. $N(0,1)$, where

i.i.d. ------------ independent, identically distributed;

$N(0,1)$ -------- normal distributions with mean $= 0$ and var $= 1$.

Denote $[n] = \{1, 2, \cdots, n\}$. Then $(X_k)_{k\in[n]}$ is a stochastic process with index set : $[n]$,

and state space : $\mathbb{R}$.

Prepared by Kelvin Shuangjian ZHANG (ks3zhang@waterloo.ca)

Focus of this term: $\mathcal{X}$ is finite or countable.
i.e., discrete state space

$T = \mathbb{N}$, i.e., discrete time

That is, discrete state space, discrete time stochastic processes.

Ex 2. $X_i \sim$ i.i.d. $\{\pm 1\}$-valued $\mathrm{Ber}(\frac{1}{2})$.

i.e. $X_i = \begin{cases} 1 & \text{w.p. } \frac{1}{2} \\ -1 & \text{w.p. } \frac{1}{2}. \end{cases}$

$(X_i)_{i=1}^{\infty}$ is a stochastic process with

index set : $\{1, 2, \cdots\}$,

and state space : $\{\pm 1\}$.

Ex 3. $(X_i)_{i=1}^{\infty}$ as in Ex 2, let $X_0 = 0$.

Define $S_n := \sum_{i=0}^{n} X_i$

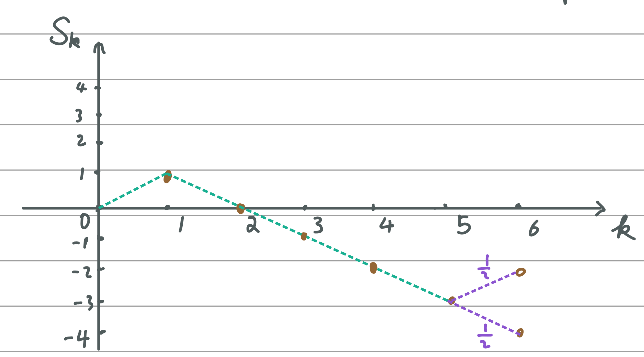
$(S_k)_{k=0}^{\infty}$ is a discrete time stochastic process with

index set : $\mathbb{N}$,

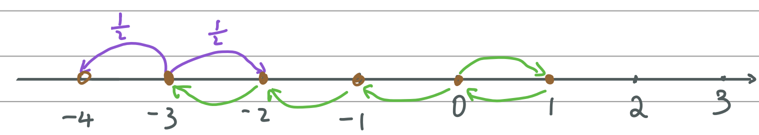
and state space : $\mathbb{Z} = \{0, \pm 1, \pm 2, \cdots\}$.

Prepared by Kelvin Shuangjian ZHANG (ks3zhang@waterloo.ca)

One can view $(S_k)_{k=0}^{\infty}$ as a path :



That is, $(S_k)_{k=0}^{\infty}$ is a simple random walk on $\mathbb{Z}$.



Part 2°. Q: Given $(X_s)_{0\le s\le t}$, how to predict $X_{t+1} = ?$

Transition probability :

$$\mathbb{P}(X_{t+1} = x_{t+1} \mid X_t = x_t, X_{t-1} = x_{t-1}, \cdots, X_0 = x_0)$$

conditional probability

can also write like: $\mathbb{P}(\{X_{t+1} = x_{t+1}\} \mid (X_s)_{0\le s\le t} = (x_s)_{0\le s\le t})$

Recall : ① Let $A, B \subseteq \Omega$ be two events s.t. $\mathbb{P}(B)>0$, let $\mathbb{P}(A \mid B) := \frac{\mathbb{P}(A,B)}{\mathbb{P}(B)}$ be the conditional probability.

(In this class, we use $\mathbb{P}(A,B)$ to denote the joint probability of A and B.)

then $\boxed{\mathbb{P}(A,B,C) = \mathbb{P}(A \mid B,C) \cdot \mathbb{P}(B \mid C) \cdot \mathbb{P}(C)}$.

a). Meaning ?

b) Can be extended to $\mathbb{P}(A_0 \mid A_1, A_2, \cdots, A_n)$, for $n > 2$.

8. Law of $(X_s)_{s \le t}$ ---------- the joint distribution

Law of X ($\mathcal{L}X : \mathcal{B} \longrightarrow \mathbb{R}$) ------------ distribution on $\mathcal{X}$.

where $X : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{X}, \mathcal{B})$ is a random variable,

$\forall S \in \mathcal{B}$, $\mathcal{L}X(S) = \mathbb{P}(X^{-1}(S))$.

Observation:

$\mathbb{P}(X_t = x_t, X_{t-1} = x_{t-1}, \cdots, X_0 = x_0)$ --------------- Joint distribution

$$
\left.
\begin{aligned}
&= \mathbb{P}(X_t = x_t \mid X_{t-1} = x_{t-1}, \cdots, X_0 = x_0) \\
&\quad \cdot \mathbb{P}(X_{t-1} = x_{t-1} \mid X_{t-2} = x_{t-2}, \cdots, X_0 = x_0) \\
&\quad \cdot \cdots \\
&\quad \cdot \mathbb{P}(X_1 = x_1 \mid X_0 = x_0)
\end{aligned}
\right] \text{------- Transition probabilities}
$$

$\cdot\ \mathbb{P}(X_0 = x_0)$ ---------------------------- Initial Data

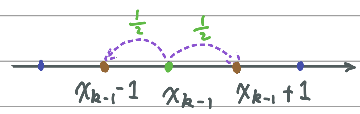
i.e. if we know transition probabilities and $\mathbb{P}(X_0 = x_0)$,

then we know the joint distribution, which means

we know the law of $(X_s)_{s \le t}$.

Ex 4. Simple Random Walk (SRW) on $\mathbb{Z}$: $(S_k)_{k\geq 0}$ has transition probabilities

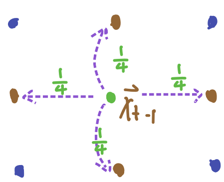
$$\mathbb{P}(S_k = x_k \mid (S_n)_{n\leq k-1} = (x_n)_{n\leq k-1}) = \begin{cases} \frac{1}{2}, & |x_k - x_{k-1}| = 1, \\ 0, & \text{else}. \end{cases}$$



S_k only depends on the previous step.

Ex 5. (SRW on $\mathbb{Z}^2$). In this case, $\mathcal{X} = \mathbb{Z}^2$, $T = \mathbb{N}$.

The transition probability



$$\mathbb{P}(X_t = (x_t^1, x_t^2) \mid X_{t-1} = \vec{x}_{t-1}, \cdots, X_0 = \vec{x}_0)$$

$$= \begin{cases} \frac{1}{4}, & \|\vec{x}_t - \vec{x}_{t-1}\| = 1; \\ 0, & \text{otherwise}. \end{cases}$$

Ex 6. Forgetful Fibonacci

Fibonacci (Italian Mathematician 1175-1250) : 1, 1, 2, 3, 5, 8, 13, 21, 34, $\cdots$

Let $X_0 \sim \text{Uniform}([n])$, $X_1 = x_0$, define the transition probability

$$\mathbb{P}(X_t = x_t \mid (X_s)_{s<t} = (x_s)_{s<t}) = \begin{cases} \frac{9}{10}, & x_t = x_{t-1} + x_{t-2}; \\ \frac{1}{10}, & x_t = x_0; \\ 0, & \text{else}. \end{cases}$$

Ex 7. (Gambler's Ruin) $(X_t)_{t\geq 0}$, $T = \mathbb{N}$.

X_t: wallet @ time t, when $X_t = 1, 2, \cdots, N-1$,

Flip a coin with probability p face up ---------- add 1 to wallet

$1-p$ face down ---------- subtract 1 from wallet.

Game stops when X_t hits 0 or N.

①. If $0 < x_{t-1} < N$,

$$\mathbb{P}(X_t = x_t \mid (X_s)_{s<t} = (x_s)_{s<t}) = \begin{cases} p, & x_t = x_{t-1} + 1; \\ 1-p, & x_t = x_{t-1} - 1; \\ 0, & \text{else}. \end{cases}$$

②. If $x_{t-1} = 0$ or N,

$$\mathbb{P}(X_t = x_t \mid (X_s)_{s<t} = (x_s)_{s<t}) = \begin{cases} 1, & x_t = x_{t-1}; \\ 0, & \text{else}. \end{cases}$$

Here, Ex 3, 4, 5, 7 are all called Markov chains.

Part 3°. Def. A stochastic process $(X_t)_{t\in T}$ is a (discrete time) Markov chain if it satisfies the Markov property:

$$\boxed{\mathbb{P}(X_t = x_t \mid (X_s)_{s<t} = (x_s)_{s<t}) = \mathbb{P}(X_t = x_t \mid X_{t-1} = x_{t-1}), \quad \forall t \geq 1, \ \forall x_0, \cdots, x_t \in \mathcal{X}.}$$

Prepared by Kelvin Shuangjian ZHANG (ks3zhang@waterloo.ca)

Ex 3, 4, 5, 7 are also time homogeneous.

Def. A Markov chain $(X_t)_{t\in T}$ is time homogeneous if

$$\boxed{\mathbb{P}(X_{t+1}=x \mid X_t = y) = \mathbb{P}(X_1 = x \mid X_0 = y), \quad \forall\, t \geq 1,\ \forall\, x, y \in \mathcal{X}.}$$

Def. For a (discrete time) time homogeneous Markov chain $(X_t)_{t=0}^{\infty}$, we call P its transition matrix if

$$\boxed{P_{xy} = \mathbb{P}(X_1 = y \mid X_0 = x), \quad \forall\, x, y \in \mathcal{X}.}$$

Here, P_{xy} represents the element of P on row x & column y, it is also called matrix P evaluated at $(x,y)\in\mathcal{X}$.

Properties of Transition Matrices:

①. $P_{xy} \in [0,1], \quad \forall (x,y)\in\mathcal{X}^2$.

②. $\forall\, x\in\mathcal{X}, \quad \sum_{y\in\mathcal{X}} P_{xy} = \sum_{y\in\mathcal{X}} \mathbb{P}(X_1 = y \mid X_0 = x)$

$$= \mathbb{P}(X_1 \in \Omega \mid X_0 = x)$$

$$= 1.$$

Prepared by Kelvin Shuangjian ZHANG (ks3zhang@waterloo.ca)

Def. A stochastic matrix is a square matrix that satisfies Property ① and ②. If, in addition, it also satisfies

③. $\forall y \in \mathcal{X}$, $\sum_{x \in \mathcal{X}} P_{xy} = 1$,

then it is called bistochastic (or doubly stochastic).

Rmk. {Stochastic matrices}

$\updownarrow$ one-to-one correspondence

{time homogeneous Markov chains}

This is the end of this lecture!