Markov Chain

Def: $\mathbb{P}(X_n = i_n \mid X_1 = i_1, \dots, X_{n-1} = i_{n-1})$

$$= \mathbb{P}(X_n = i_n \mid X_{n-1} = i_{n-1})$$

knowing present, past and future independent

e.g. (6.1.1)

Any seq of indep r.v. $\{X_n\}$ that take values in countable set S is a Markov chain.

Pf: $\forall\, i_1, \dots, i_n \in S$,

$$\mathbb{P}(X_n = i_n \mid X_1 = i_1, \dots, X_{n-1} = i_{n-1})$$
$$\parallel$$
$$\mathbb{P}(X_n = i_n)$$
$$\parallel$$
$$\mathbb{P}(X_n = i_n \mid X_{n-1} = i_{n-1})$$

In particular, time-homogeneous iff $\{X_n\}$ i.i.d.

since $\mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(X_{n+1} = j)$ does not depend on n iff $\{X_n\}$ i.i.d.

e.g. (6.1.2)

A dice rolled repeatedly, which are Markov?

(a): $X_n \triangleq$ largest num shown up to the nth roll

Pf: denote $\xi_1, \dots, \xi_n$ as the outcome of n rolls

$X_n = \max\{\xi_1, \dots, \xi_n\} = \max\{X_{n-1}, \xi_n\}$

$$\mathbb{P}(X_n = i_n \mid X_1 = i_1, \dots, X_{n-1} = i_{n-1})$$

$$= \mathbb{P}(\max\{i_{n-1}, \xi_n\} = i_n \mid X_1 = i_1, \dots, X_{n-1} = i_{n-1})$$

indep

$$= \mathbb{P}(\max\{i_{n-1}, \xi_n\} = i_n) \overset{\text{similarly}}{=} \mathbb{P}(X_n = i_n \mid X_{n-1} = i_{n-1})$$

✓

	1	2	3	4	5	6
1	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
2		$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
3			$\frac{3}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
4				$\frac{4}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
5					$\frac{5}{6}$	$\frac{1}{6}$
6						1

(b): $N_n \triangleq$ num of sixes in n rolls

Pf: $N_n = \sum_{i=1}^{n} I_{\{\xi_i = 6\}} = \underline{N_{n-1} + I_{\{\xi_n = 6\}}}$ random walk structure

$\mathbb{P}(N_n = i_n \mid N_1 = i_1, \dots, N_{n-1} = i_{n-1})$

$= \mathbb{P}(i_{n-1} + I_{\{\xi_n = 6\}} = i_n \mid N_1 = i_1, \dots, N_{n-1} = i_{n-1})$

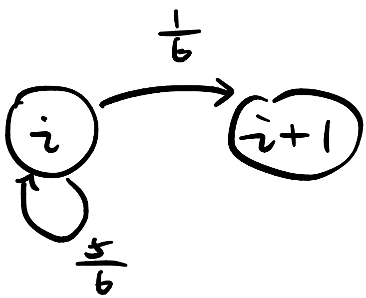
indep.

$= \mathbb{P}(i_{n-1} + I_{\{\xi_n = 6\}} = i_n) = \mathbb{P}(N_n = i_n \mid N_{n-1} = i_{n-1})$ ✓

similarly

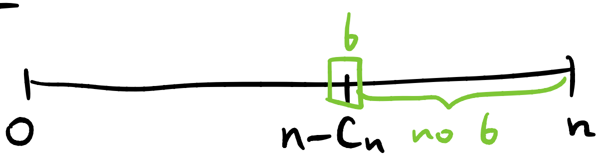
RW is Markov!

state space $S = \mathbb{N}$



(c): $C_n \triangleq$ the time since the most recent six at time n

Pf:



$$C_n = \begin{cases} C_{n-1} + 1 & \text{if } \xi_n \neq 6 \\ 0 & \text{if } \xi_n = 6 \end{cases} = (C_{n-1} + 1) \cdot I_{\{\xi_n \neq 6\}}$$

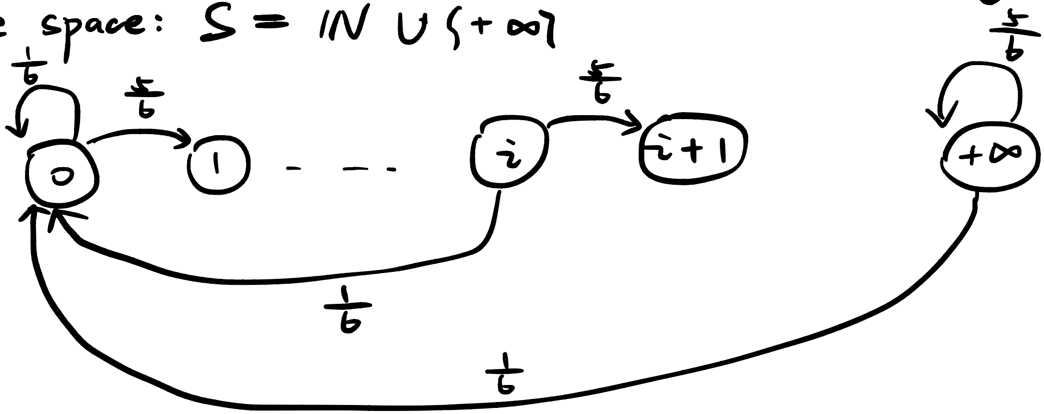
$$\mathbb{P}(C_n = i_n \mid C_1 = i_1, \dots, C_{n-1} = i_{n-1})$$

$$= \mathbb{P}\big((i_{n-1}+1)\cdot I_{\{f_n \neq 6\}} = i_n \mid C_1 = i_1, \dots, C_{n-1} = i_{n-1}\big)$$

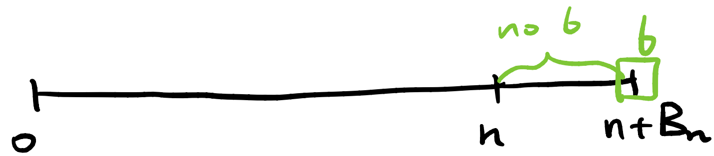
indep

$$= \mathbb{P}\big((i_{n-1}+1)\cdot I_{\{f_n \neq 6\}} = i_n\big) = \mathbb{P}(C_n = i_n \mid C_{n-1} = i_{n-1})$$

State space: $S = \mathbb{N} \cup \{+\infty\}$



(d): $B_n \triangleq$ the time until the next 6 at time n



$$B_n = \begin{cases} B_{n-1} - 1 & \text{if } B_{n-1} > 0 \\ \underline{G_n} & \text{if } B_{n-1} = 0 \end{cases}$$

a r.v. $\sim G(\frac{1}{6})$ indep of $\{B_n\}$

$$= (B_{n-1} - 1)\cdot I_{\{B_{n-1} > 0\}} + G_n \cdot I_{\{B_{n-1} = 0\}}$$

So $\mathbb{P}(B_n = i_n \mid B_1 = i_1, \ldots, B_{n-1} = i_{n-1})$

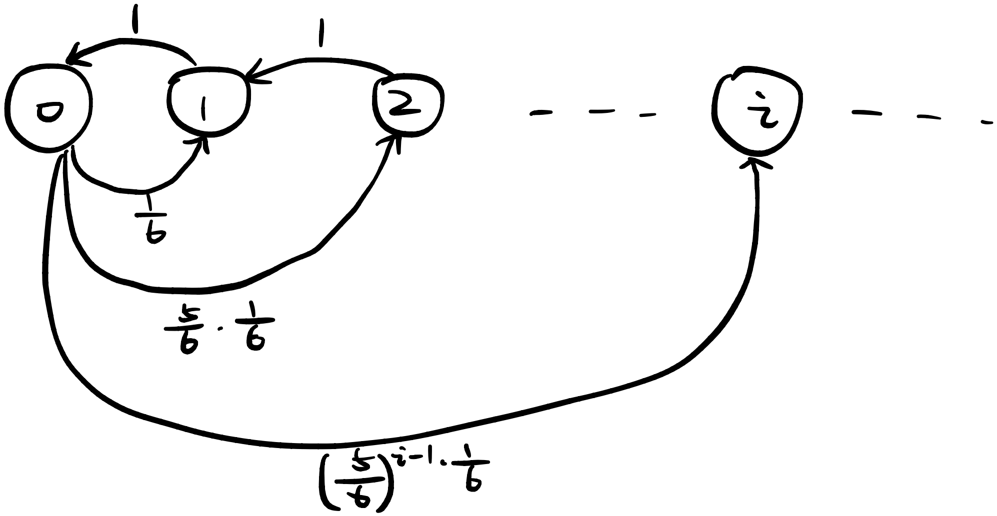
$$= \mathbb{P}\left((i_{n-1}-1)\cdot I_{\{i_{n-1}>0\}} + \underline{\underline{G_n}} \cdot I_{\{i_{n-1}=0\}} \,\middle|\, B_1 = i_1, \ldots, B_{n-1} = i_{n-1} \right)$$

Markov

$$= \mathbb{P}\left((i_{n-1}-1) I_{\{i_{n-1}>0\}} + G_n \cdot I_{\{i_{n-1}=0\}} \right)$$

$$= \mathbb{P}(B_n = i_n \mid B_{n-1} = i_{n-1}) \quad \checkmark$$

State space: $S = \mathbb{N}$



Why $+\infty \notin S$?

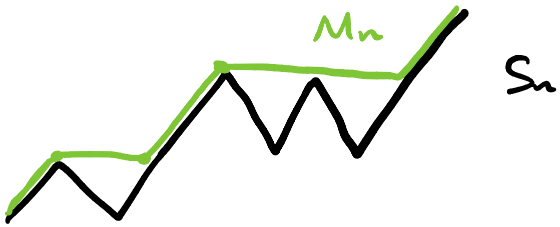
$\mathbb{P}(6 \text{ never appears in } \xi_1, \ldots, \xi_n) = \left(\frac{5}{6}\right)^n \to 0 \quad (n \to \infty)$

So with probability 1, we see at least one 6.

eg: (6.1.3) $\{S_n\}$ SRW, $S_0 = 0$, $M_n = \sup_{k \le n} \{S_k\}$,

show $Y_n = M_n - S_n$ is Markov.

Pf:



Y_n is the gap in between

$$S_n = S_{n-1} + \boxed{\xi_n} \quad \text{i.i.d. movements}$$

$$Y_n = \begin{cases} Y_{n-1} - 1 & \text{if } Y_{n-1} > 0, \ \xi_n = 1 \\ Y_{n-1} + 1 & \text{if } Y_{n-1} > 0, \ \xi_n = -1 \\ 0 & \text{if } Y_{n-1} = 0, \ \xi_n = 1 \\ 1 & \text{if } Y_{n-1} = 0, \ \xi_n = -1 \end{cases}$$

$$= (Y_{n-1} - \xi_n) \cdot I_{\{Y_{n-1} > 0\}} + I_{\{Y_{n-1} = 0, \xi_n = -1\}}$$

So $\mathbb{P}(Y_n = i_n \mid Y_1 = i_1, \dots, Y_{n-1} = i_{n-1})$

$$= \mathbb{P}\left((i_{n-1} - \xi_n) \cdot I_{\{i_{n-1} > 0\}} + I_{\{i_{n-1} = 0, \xi_n = -1\}} = i_n \;\middle|\; \begin{matrix} Y_1 = i_1, \\ \dots, \\ Y_{n-1} = i_{n-1} \end{matrix} \right)$$

indep

$= \dots$

$\overset{\text{similar}}{=} \mathbb{P}(Y_n = i_n \mid Y_{n-1} = i_{n-1})$ ✓

e.g.: (6.1.6)

T is stopping time ($I_{\{T=n\}}$ is func of $X_0, \dots, X_n$),

then if $\{X_n\}$ is Markov chain (time-homogeneous), then

strong Markov property holds:

$$\mathbb{P}(X_{T+m}=j \mid X_0=x_0, \dots, X_T=x_T) = \mathbb{P}(X_{T+m}=j \mid X_T=x_T)$$

Pf: Turn random time to deterministic time:

$$\mathbb{P}(X_{T+m}=j \mid X_0=x_0, \dots, X_T=x_T)$$

$$= \frac{\mathbb{P}(X_0=x_0, \dots, X_T=x_T, X_{T+m}=j)}{\mathbb{P}(X_0=x_0, \dots, X_T=x_T)}$$

$$= \frac{\sum_{k=0}^{\infty} \mathbb{P}(X_0=x_0, \dots, X_T=x_T, X_{T+m}=j, T=k)}{\mathbb{P}(X_0=x_0, \dots, X_T=x_T)}$$

$$= \frac{\sum_{k=0}^{\infty} \mathbb{P}(X_0=x_0, \dots, X_k=x_T, X_{k+m}=j, T=k)}{\mathbb{P}(X_0=x_0, \dots, X_T=x_T)}$$

$\mathbb{P}(X_{k+m}=j \mid X_k=x_T)$ by Markov

$$= \frac{\sum_{k=0}^{\infty} \mathbb{P}(X_{k+m}=j \mid X_0=x_0, \dots, X_k=x_T, T=k) \cdot \mathbb{P}(X_0=x_0, \dots, X_T=x_T, T=k)}{\mathbb{P}(X_0=x_0, \dots, X_T=x_T)}$$

time-homogeneous

$$= \frac{P_{x_T,j}(m) \cdot \mathbb{P}(X_0=x_0, \cancel{\dots, X_T=x_T})}{\mathbb{P}(X_0=x_0, \cancel{\dots, X_T=x_T})} = P_{x_T,j}(m)$$

Similarly, $\mathbb{P}(X_{T+m}=j \mid X_T=x_T) = P_{x_T,j}(m)$ ✓

☆: since X_k is given, $\{T=k\}$ actually has no dependence on X_k!

Why does $\mathbb{P}(X_{k+m}=j \mid X_0=x_0, \dots, X_k=x_T, T=k)$

$$= \mathbb{P}(X_{k+m}=j \mid X_k=x_T) \; ?$$

Since $\underline{X_{k+m}\big|_{f(X_0,\dots,X_k),\, X_k} \overset{d}{=} X_{k+m}\big|_{X_k} \text{ for } \forall f}$

Pf:

$$\mathbb{P}(X_{k+m}=x_{k+m} \mid f(X_0,\dots,X_k)=a,\; X_k=x_k)$$

$$= \mathbb{P}(X_{k+m}=x_{k+m} \mid f(X_0,\dots,X_{k-1},x_k)=a,\; X_k=x_k)$$

$$= \frac{\mathbb{P}(X_{k+m}=x_{k+m},\, X_k=x_k,\, f(X_0,\dots,X_{k-1},x_k)=a)}{\mathbb{P}(f(X_0,\dots,X_{k-1},x_k)=a,\; X_k=x_k)}$$

$$= \frac{\sum_{x_0,\dots,x_{k-1}} \mathbb{P}(X_{k+m}=x_{k+m},\, X_k=x_k,\, X_0=x_0,\dots,X_{k-1}=x_{k-1})}{\mathbb{P}(f(\dots),\, X_k)}$$

all values of $(x_0,\dots,x_{k-1})$ s.t. $f(x_0,\dots,x_{k-1},x_k)=a$.

$$\overset{\text{Markov}}{=} \frac{\sum_{x_0,\dots,x_{k-1}} \mathbb{P}(X_{k+m}=x_{k+m} \mid X_k=x_k)\cdot \mathbb{P}(X_0=x_0,\dots,X_k=x_k)}{\mathbb{P}(f(\dots),\, X_k)}$$

$$= \mathbb{P}(X_{k+m}=x_{k+m} \mid X_k=x_k) \quad ✓$$