Symmetry Ideas: General idea of symmetry: make a transformation, and something stays the same. In probability theory, the transformation may be conditioning, or some rearrangement of variables. What stays the same is the distribution of something. For a sequence of random variables $X_0, X_1, X_2, \ldots$, various notions of symmetry:

1. Independence. $X_0$ and $X_1$ are independent. Distribution of $X_1$ given $X_0$ does not involve $X_0$; that is, $(X_1|X_0 \in A) \overset{d}{=} X_1$.

2. Identical distribution. The $X_n$ are identically distributed: $X_0 \overset{d}{=} X_n$ for every $n$. Of course IID $\implies$ LLN / CLT.

Stationary: $(X_1, X_2, \ldots) \overset{d}{=} (X_0, X_1, \ldots)$ (shifting time by 1)

By measure theory, this is the same as

$$(X_1, X_2, \ldots, X_n) \overset{d}{=} (X_0, X_1, \ldots, X_{n-1}), \text{ for all } n.$$ 

Obviously IID $\implies$ Stationary. But the converse is not necessarily true. E.g., $X_n = X_0$ for all $n$, for any non-constant random variable $X_0$.

Another example: Stationary MC, i.e., start a MC with transition matrix $P$ with initial distribution $\pi$. Then easily, the following are equivalent:

$$\pi P = \pi$$

$$X_1 \overset{d}{=} X_0$$

$$(X_1, X_2, \ldots) \overset{d}{=} (X_0, X_1, \ldots)$$

Otherwise put, $(X_n)$ is stationary means $(X_1, X_2, \ldots)$ is a Markov chain with exactly the same distribution as $(X_0, X_1, \ldots)$.

Other symmetries:

- Cyclic symmetry: $(X_1, X_2, \ldots, X_n) \overset{d}{=} (X_2, X_3, \ldots, X_n, X_1)\overset{n}{\quad}$

- Reversible: $(X_n, X_{n-1}, \ldots, X_1) \overset{d}{=} (X_1, X_2, \ldots, X_n)$
Lecture 19: Stationary Markov Chains

- Exchangeable: \((X_{\pi(1)}, X_{\pi(2)}, \ldots, X_{\pi(n)}) \overset{d}{=} (X_1, X_2, \ldots, X_n)\) for every permutation \(\pi\) of \((1, \ldots, n)\).

- Application of stationary idea to MC’s.

Think about a sequence of RVs \(X_0, X_1, \ldots\) which is stationary. Take a set \(A\) in the state space. Let \(T_A = \left\{ \text{least } n \geq 1 \text{ (if any) s.t. } X_n \in A \right\}\); that is, \(T_A := \text{least } n \geq 1 : 1_A(X_n) = 1\). Consider the \(1_A(X_n)\) process.

E.g., \(0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ldots\)

To provide an informal notation:

\[
\begin{align*}
\mathbb{P}(T_A = n) &= \mathbb{P}(\underbrace{0 \ 0 \ldots \ 0 \ 0 \ 1}_{\text{n-1 zeros}}) \\
\mathbb{P}(T_A \geq n) &= \mathbb{P}(\underbrace{0 \ 0 \ldots \ 0 \ 0}_{\text{n-1 zeros}} \ ?) \\
\mathbb{P}(X_0 \in A, T_A \geq n) &= \mathbb{P}(\underbrace{1 \ 0 \ 0 \ldots \ 0 \ 0}_{\text{n-1 zeros}} \ ?)
\end{align*}
\]

Lemma: For every stationary sequence \(X_0, X_1, \ldots,\)

\[
\mathbb{P}(T_A = n) = \mathbb{P}(X_0 \in A, T_A \geq n) \quad \text{for } n = 1, 2, 3, \ldots
\]

With the informal notation, the claim is:

\[
\mathbb{P}(\underbrace{0 \ 0 \ldots \ 0 \ 1}_{\text{n-1 zeros}}) = \mathbb{P}(\underbrace{1 \ 0 \ 0 \ldots \ 0}_{\text{n-1 zeros}} \ ?)
\]

Notice. This looks like reversibility, but reversibility of the sequence \(X_0, X_1, \ldots, X_n\) is not being assumed. Only stationarity is required!

Proof
Take the identity above, sum over \( n = 1, 2, 3, \ldots \),

\[
P(T_A < \infty) = \sum_{n=1}^{\infty} P_A(T = n)
\]

\[
= \sum_{n=1}^{\infty} P(X_0 \in A, T_A \geq n)
\]

\[
= E \sum_{n=1}^{\infty} 1(X_0 \in A, T_A \geq n)
\]

\[
= E(T_A 1(X_0 \in A))
\]

This is Marc Kac’s identity: For every stationary sequence \((X_n)\), and every measurable subset \(A\) of the state space of \((X_n)\):

\[
P(T_A < \infty) = E(T_A 1(X_0 \in A))
\]

- Application to recurrence of Markov chains: Suppose that an irreducible transition matrix \(P\) on a countable space \(S\) has a stationary probability measure \(\pi\), that is \(\pi P = \pi\), with \(\pi_i > 0\) for some state \(i\). Then
  - \(\pi\) is the unique stationary distribution for \(P\)
  - \(\pi_j > 0, \forall j\).
  - \(E_j T_j < \infty, \forall j\).
  - \(E_j T_j = \frac{1}{\pi_j}, \forall j\).

Proof:

Apply Kac’s identity to \(A = \{i\}\), with \(P = P_\pi\) governing \((X_n)\) as a MC with transition matrix \(P\) started with \(X_0 \overset{\text{d}}{=} \pi\).

\[
1 \geq P_\pi(T_i < \infty) \overset{\text{Kac}}{=} E_\pi[T_i 1(X_0 = i)] = \pi_i E_i T_i
\]

Since \(\pi_i > 0\) this implies first \(E_i T_i < \infty\), and hence \(P_i(T_i < \infty) = 1\).

Next, need to argue that \(P_j(T_i < \infty) = 1\) for all states \(j\). This uses irreducibility, which gives us an \(n\) such that \(P^n(i, j) > 0\), hence easily an \(m \leq n\) such that it is possible to get from \(i\) to \(j\) in \(m\) steps without revisiting \(i\) on the way, i.e.

\[
P_i(T_i > m, X_m = j) > 0
\]

But using the Markov property this makes

\[
P_j(T_i < \infty) = P_i(T_i < \infty | T_i > m, X_m = j) = 1
\]
Finally
\[ P_\pi(T_i < \infty) = \sum_j \pi_j P_j(T_i < \infty) = \sum_j \pi_j = 1 \]

and the Kac formula becomes \( 1 = \pi_i \mathbb{E}_i T_i \) as claimed. Lastly, it is easy that \( \pi_j > 0 \) and hence \( 1 = \pi_j \mathbb{E}_j T_j \) because \( \pi = \pi P \) implies \( \pi = \pi P^n \) and so
\[ \pi_j = \sum_k \pi_k P^n(k, j) \geq \pi_i P^n(i, j) > 0 \]

for some \( n \) by irreducibility of \( P \).