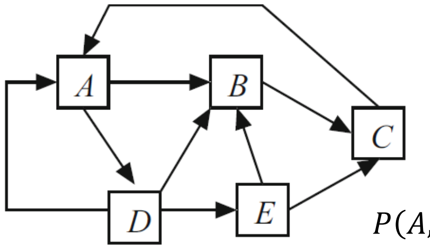
EECS 126: Probability & Random Processes
Fall 2021

PageRank

Shyam Parekh

PageRank

- Originally used by Google for ranking the pages from a keyword search.



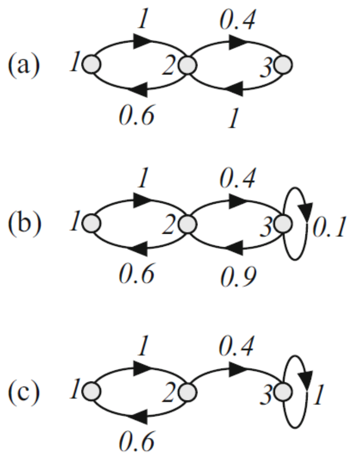
$P(A, B) = 1/2, P(D, E) = 1/3, \ldots$

$$\pi(i) = \sum_{j \in X} \pi(j) P(j, i),\ \forall\, i\ \in\ \mathcal{X}\ \Leftrightarrow \pi = \pi P$$

- Original algorithm also used a damping factor.
- With the normalization condition $\sum_{i \epsilon X} \pi(i) = 1, \pi = \frac{1}{39}[12, 9, 10, 6, 2]$
- This is similar to a Markov Chain (defined next).

Discrete Time Markov Chain (DTMC)

- DTMC $\{X(n), n \geq 0\}$ over state space $\mathcal{X}$ with $P = [P(i,j)]$ as the transition probability matrix with $P(i,j) = P[X(n+1) = j \mid X(n) = i]$.
  - State transition diagram.
  - Memoryless Property: $P[X(n+1) = j \mid X(n) = i, X(m), m < n] = P(i,j) \; \forall \; i, j$.
- $\pi_{n+1}(\mathrm{i}) = \sum_{j \in X} \pi_n(j) P(j,i) \Leftrightarrow \pi_{n+1} = \pi_n P \Leftrightarrow \pi_n = \pi_0 P^n, \forall \; n \geq 0.$
  - A normalized π is an invariant distribution $\Leftrightarrow \pi = \pi \mathrm{P}$.
  - Irreducible: MC can reach any state from any other state (possibly in multiple steps).
  - Aperiodic: Let $d(i) \coloneqq g.c.d.\{n \geq 1 \mid P^n(i,i) > 0\}$. An irreducible DTMC is aperiodic if $d(i) = 1, \forall \; i$. (Fact: In an irreducible DTMC, $d(i)$ is same ∀ i.)



Irreducible & Periodic

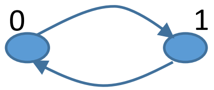
Irreducible & Aperiodic

Reducible

Big Theorem for Finite DTMC

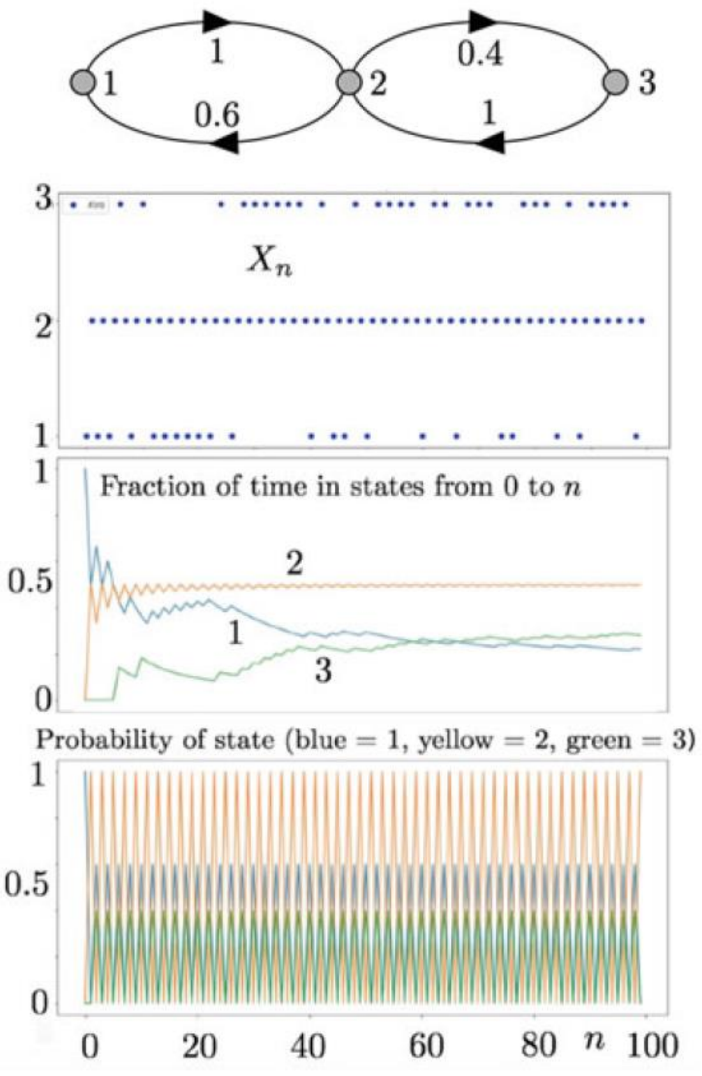
- Theorem: Consider an irreducible DTMC over a finite state space. Then,
  (a) There is a unique invariant distribution π.
  (b) Long-term fraction of time $(X(n) = i) := \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} 1\{X(n) = i\} = \pi(i)$.
  (c) If the DTMC is aperiodic, $\pi_n \to \pi$.

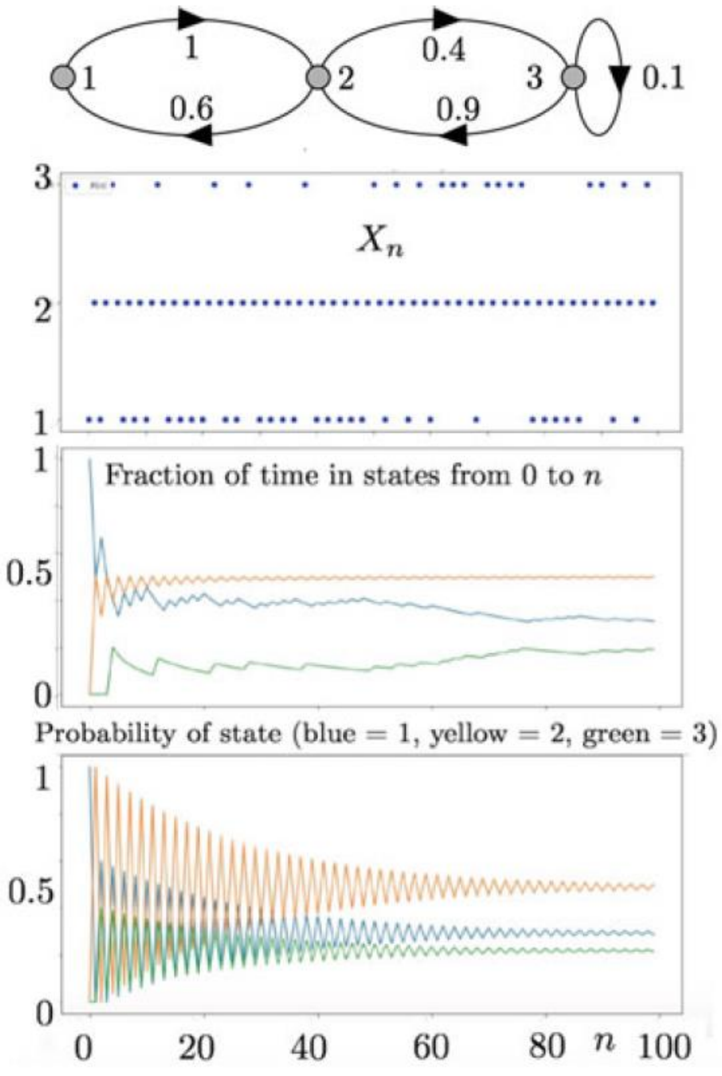
- Example showing aperiodicity is necessary for (c).

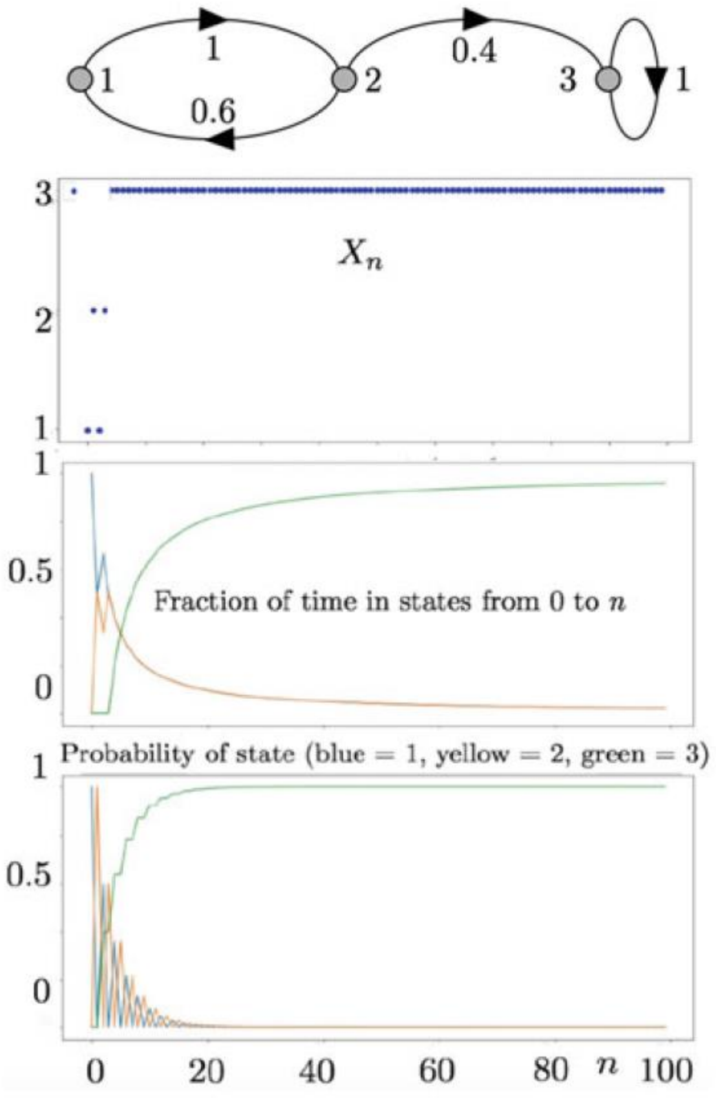


- In part (c), $\pi_n \to \pi$ irrespective of the initial distribution π_0.
  - This implies each row of P^n converges to π as $n \to \infty$.

Illustrations

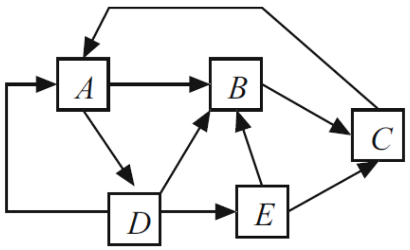






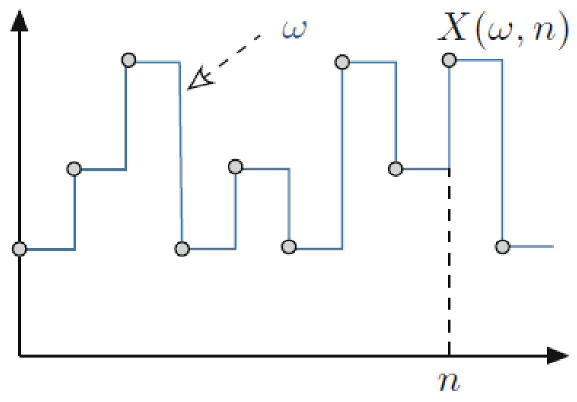
Hitting Times

1. Mean Hitting Time $\beta(i) := E[T_A \mid X_0 = i],\ i \in \mathcal{X},\ A \subset \mathcal{X}$.
   - First Step Equations (FSEs) are: $\beta(i) = \begin{cases} 1 + \sum_j P(i,j)\beta(j), if\ i \notin A \\ 0, if\ i \in A \end{cases}$
2. Probability of hitting a state before another $\alpha(i) := P[T_A < T_B \mid X_0 = i]$, $i \in \mathcal{X},\ A, B \subset \mathcal{X},\ A \cap B = \emptyset$.
   - FSEs are: $\alpha(i) = \begin{cases} \sum_j P(i,j)\alpha(j), if\ i \notin A \cup B \\ 1, if\ i \in A \\ 0, if\ i \in B \end{cases}$
3. Average discounted reward $\delta(i) := E[Z \mid X_0 = i], i \in \mathcal{X}$, where $Z = \sum_{n=0}^{T_A} \beta^n h(X(n)), A \subset \mathcal{X}$ with $0 < \beta \leq 1$ as the discount factor.
   - FSEs are: $\delta(i) = \begin{cases} h(i) + \beta \sum_j P(i,j)\delta(j), if\ i \notin A \\ h(i), if\ i \in A \end{cases}$



Markov Chain Trajectory/Realization

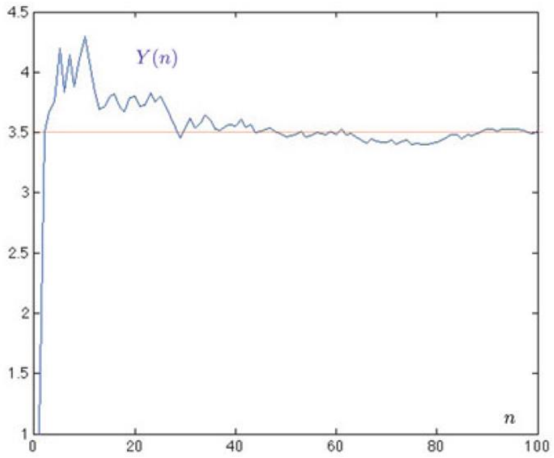
- Canonical Sample Space with outcome ω being the Markov Chain trajectory/realization:



- $P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \pi(i_0)P(i_0, i_1) \ldots P(i_{n-1}, i_n)$
- There exists a consistent probability measure over the sample space.

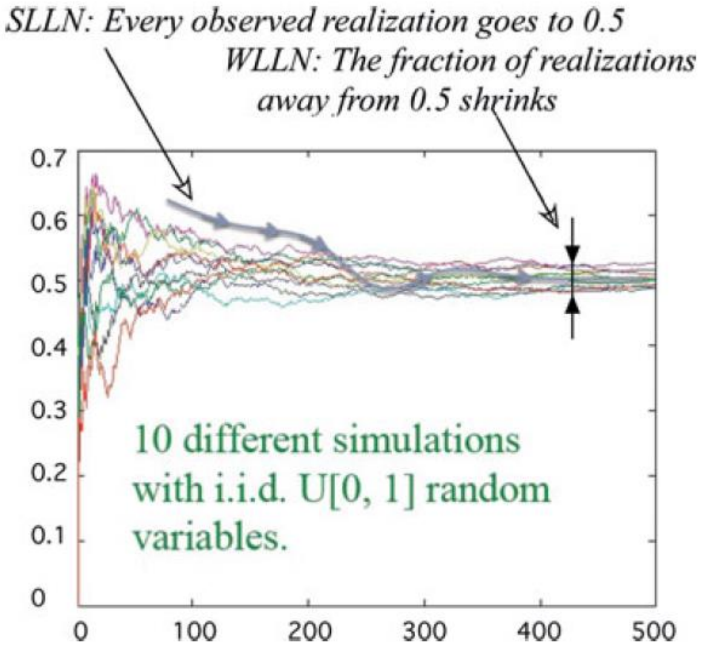
Laws of Large Numbers

- Let $\{X(i), i \geq 1\}$ be a sequence of Independent and Identically Distributed (IID) RVs with mean μ, and let $S(n) = \sum_{i=1}^{n} X(i)$. Assume $E(|X(i)|) < \infty$.
- Strong Law of Large Numbers (SLLN): $P(\lim_{n\to\infty} \frac{S(n)}{n} = \mu) = 1$.
  - $\frac{S(n)}{n}$ converges to μ almost surely.
- Weak Law of Large Numbers (WLLN): $\lim_{n\to\infty} P(\left|\frac{S(n)}{n} - \mu\right| > \epsilon) = 0, \forall\, \epsilon > 0$.
  - $\frac{S(n)}{n}$ converges to μ in probability.
- Almost sure convergence implies convergence in probability.
- SLLN Example: $Y(n) = \frac{S(n)}{n}$, where $X(i)$ = outcome of ith roll of a balanced die.



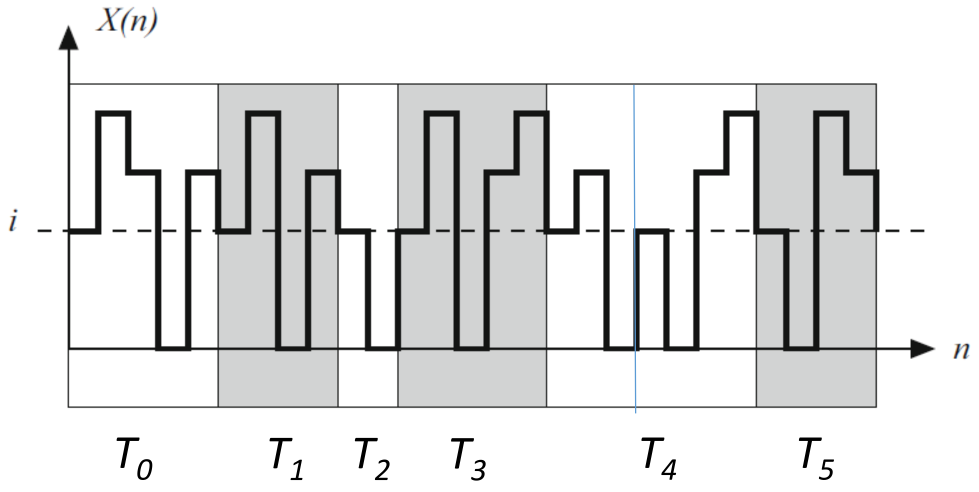
- Part (b) of the Big Theorem indicates almost sure convergence.
  - For an irreducible DTMC over a finite numerical state space, $\frac{1}{n}\sum_{i=0}^{n-1} X(i)$ converges almost surely to $E(X)$ where X is an RV with the distribution same as the invariant distribution π.

SLLN vs. WLLN, SLLN Proof



- Let $\{X(i), i \geq 1\}$ be a sequence of Independent and Identically Distributed (IID) RVs with mean μ, and let $S(n) = \sum_{i=1}^{n} X(i)$. Assume $E(X(i)^4) < \infty$.
- Strong Law of Large Numbers (SLLN): $P(\lim_{n\to\infty} \frac{S(n)}{n} = \mu) = 1$.

Key Points from the Big Theorem Proof



- T_i = Time to return to *i*. T_i's are IID.