1. The state space $\mathcal{S}$ is either $\{1, 2, \ldots, M\}$ for some integer $M > 0$ or $\{1, 2, 3, 5, \ldots\}$.

2. The state of the Markov chain at time $n \in \mathbb{N}$ is the $\mathcal{S}$-valued random quantity X_n.

3. For any states i and j in $\mathcal{S}$,

$$P(i,j) = \text{Prob}\{X_{n+1} = j \mid X_n = i\} \ ,$$

and this one-step transition probability doesn't depend on n because we assume the Markov chain is homogeneous. The same is true of the m-step transition probabilities

$$P^{(m)}(i,j) = \text{Prob}\{X_{n+m} = j \mid X_n = i\} \ ,$$

which you can calculate inductively from the $P(i,j)$ via the recursion

$$P^{(m+1)}(i,j) = \sum_{k \in \mathcal{S}} P(i,k) P^{(m)}(k,j) \ .$$

4. The initial distribution $\pi(0)$ is a list of nonnegative numbers summing to 1, one for each state in $\mathcal{S}$. For any $j \in \mathcal{S}$,

$$\pi_j(0) = \text{Prob}\{X_0 = j\} \ .$$

5. We use Prob_i and E_i notation a lot. For any $i \in \mathcal{S}$,

$$\text{Prob}_i\{\text{thing}\} = \text{Prob}\{\text{thing} \mid X_0 = i\}$$

and

$$E_i(\text{thing}) = E(\text{thing} \mid X_0 = i) \ .$$

Observe that $P(i,j) = \text{Prob}_i\{X_1 = j\}$. Observe also that, by the rules of conditional probability,

$$\text{Prob}\{\text{thing}\} = \sum_{i \in \mathcal{S}} \pi_i(0) \text{Prob}_i\{\text{thing}\}$$

and

$$E(\text{thing}) = \sum_{i \in \mathcal{S}} \pi_i(0) E_i(\text{thing}) \ .$$

6. First hitting times are important random quantities associated with the Markov chain. For each $j \in \mathcal{S}$,

$$T_j = \text{the first time } n > 0 \text{ when } X_n = j \ .$$

Note that T_j is never zero, and $T_j = \infty$ if and only if $X_n \neq j$ for every $n > 0$. In other words, $T_j = \infty$ precisely when the Markov chain follows a path that never hits j after time $n = 0$.

7. For each $j \in \mathcal{S}$, another important random quantity, defined for every $n > 0$, is

$$N_j(n) = \text{the number of times } X_m = j \text{ during the interval } 1 \leq m \leq n \ .$$

Note that $0 \leq N_j(n) \leq n$ for all j and n and that $N_j(n)$ is weakly increasing in n. For each $j \in \mathcal{S}$, set

$$N_j = \lim_{n \to \infty} N_j(n) \ .$$

N_j represents the total number of times $n > 0$ for which $X_n = j$, and $N_j = \infty$ is allowed.

8. For any pair of states i and j,

$$f_{ij}^{(k)} = \text{Prob}_i\{k \text{ is the smallest } n > 0 \text{ for which } X_n = j\}$$

Alternatively,

$$f_{ij}^{(k)} = \text{Prob}_i \{T_j = k\} .$$

Also set

$$r_{ij} = \text{Prob}_i\{X_n = j \text{ for some (finite) } n > 0\} .$$

Note that

$$r_{ij} = \sum_{k=1}^{\infty} f_{ij}^{(k)} .$$

9. For any pair of states i and j, $i \to j$ means there's a positive probability that the Markov chain reaches state j at some finite positive time given that it starts in state i. Note that $i \to j$ is the same as $r_{ij} > 0$. It's also the same as $f_{ij}^{(k)} > 0$ for some $k > 0$. It's also the same as $P^{(m)}(i, , j) > 0$ for some $m > 0$. It's not the same as $P(i, j) > 0$, i.e. it's not the same as saying that the transition diagram for the Markov chain has an arrow leading directly from state i to state j.

10. A state $j \in \mathcal{S}$ is transient when $r_{jj} < 1$ and recurrent when $r_{jj} = 1$.

11. For any $j \in \mathcal{S}$, set

$$m_j = E_j\left(T_j\right) = \sum_{k=1}^{\infty} k f_{jj}^{(k)} .$$

If j is transient, $m_j = \infty$. If j is recurrent, we say j is positively recurrent when $m_j < \infty$ and j is null-recurrent when $m_j = \infty$. A finite-state Markov chain has no null-recurrent states. Infinite m_j means that if you start in state j, you have to wait on average infinitely long to see state j again.

12. The state space $\mathcal{S}$ parses into disjoint sets. One is the set of all transient states. The others are closed sets of states called recurrence classes. In any recurrence class C, $i \to j$ for all i and j in C. Any recurrence class contains either only positively recurrent states or only null-recurrent states.

13. Just as $\pi(0)$ describes X_0's probability distribution, $\pi(n)$ describes X_n's probability distribution in the sense that

$$\pi_j(n) = \text{Prob}\{X_n = j\} \text{ for all } j \in \mathcal{S} .$$

You can compute $\pi(n)$ from the recursion

$$\pi_j(n+1) = \sum_{i \in \mathcal{S}} \pi_i(n) P(i, j)$$

initialized with $\pi(0)$. $\overline{\pi}$ is a stationary distribution when $\overline{\pi}$ is a fixed point of that recursion in the sense that when $\pi(0) = \overline{\pi}$ we have $\pi(n) = \overline{\pi}$ for all $n > 0$.

14. A Markov chain can have no stationary distributions, exactly one stationary distribution, or infinitely many stationary distributions. Every stationary distribution assigns probability zero to transient and null-recurrent states.

15. Among the several limit theorems we addressed, perhaps the most important applies to a special but generic family of Markov chains, those that have only positively recurrent states and are also irreducible in the sense that $i \to j$ for any pair of states i and j. The entire state space $\mathcal{S}$ of such a Markov chain constitutes one recurrence class, and for any initial distribution $\pi(0)$ we have

$$\lim_{n \to \infty} \frac{N_j(n)}{n} = \frac{1}{m_j}$$

with probability 1. Any such Markov chain has a unique stationary distribution π^* given by

$$\pi_j^* = \frac{1}{m_j} \quad \text{for all} \quad j \in \mathcal{S} \ .$$

Thus we can describe the time-evolution of the state of any such Markov chain as follows: for any initial distribution $\pi(0)$, with probability 1 the Markov chain spends in each state j a limiting fraction of the time equal to j's "weight" according to the unique stationary distribution π^*.s

16. A direct consequence of the limit theorem just quoted is the following version of the Ergodic Theorem for Markov chains. Suppose a Markov has only positively recurrent states and is also irreducible in the sense that $i \to j$ for any pair of states i and j. Let π^* be the Markov chain's unique stationary distribution and let $f : \mathcal{S} \to \mathbb{R}$ be any function with finite π^*-mean, i.e.

$$E_{\pi^*}(f) = \sum_{j \in \mathcal{S}} \pi_j^* f(j) < \infty \ .$$

For any stationary distribution $\pi(0)$, we have

$$\lim_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} f\left(X_m\right) = E_{\pi^*}(f)$$

with probability 1.