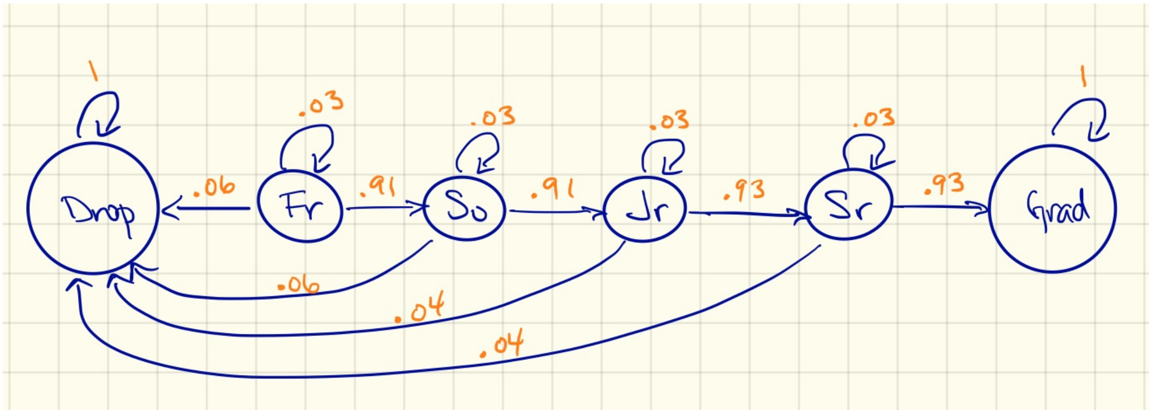
Introduction to Markov Chains

Stat 340, Fall 2021

Example.[1] University administrators have developed a Markov chain model to simulate graduation rates at their school. Students might drop out, repeat a year, or move on to the next year. Below is a graph representing the possible transitions that students can make at the university. Probabilities are listed next to each possible path, and only paths with positive probability are drawn.



Your turn:

1. What's the probability that a student who drops out will re-enroll?

2. What's the probability that a senior will graduate?

3. Does that probability depend on how many years it took them to achieve senior class standing?

Definition: Markov chain

A sequence of random variables, $X_0, X_1, X_2, \ldots$, taking values in the *state space* $\{1, \ldots, M\}$ is called a Markov chain if for all $n \geq 0$

Remarks:

- Think of X_n as the state of the system at (discrete) time n.
- q_{ij} is the *transition probability* from state i to state j.

[1] Source: Introduction to Stochastic Processes with R by Bob Dobrow

Definition: Transition matrix

A Markov chain can be represented by an $M \times M$ matrix of the probabilities $Q = (q_{ij})$

where the rows represent the

and the columns represent the

Your turn: Write down the 6×6 transition matrix for the university graduation rate Markov chain model.

4. Should the probabilities within each row sum to 1?

5. Should the probabilities within each row sum to 1?

Calculating probabilities using the transition matrix

If we know the transition matrix, Q, then we can derive the probability that a student goes from state i to state j in some given number of steps.

One-step transition probability: $P(X_{n+1} = j | X_n = i) =$

Two-step transition probability: $P(X_{n+2} = j | X_n = i) =$

m-step transition probability: $P(X_{n+m} = j | X_n = i) =$

Marginal distribution of $\mathbf{X_n}$

Suppose that at time n, X_n has PMF given by $\mathbf{s} = (s_1, s_2, \ldots, s_m)$ where $s_i = P(X_n = i)$.

We can use the law of total probability to derive the PMF of X_{n+1}:

Uses of Markov Chains

1. Use a Markov chain model, if your Markov chain is a reasonable abstraction of reality.

2. **Markov Chain Monte Carlo (MCMC)**. Synthetically construct a Markov chain that is known to converge to the distribution of interest.

Not all Markov chains will converge to a single distribution, so we need a few more concepts before we can explore MCMC.

Classification of states

Definition: A state is **recurrent** if starting there, the chain has probability 1 of returning to that state.

Definition: A state that is not recurrent is **transient**.

Definition: If it's possible to get from any state to any state in a chain (with positive probability) *in a finite number of steps*, then it is **irreducible**.

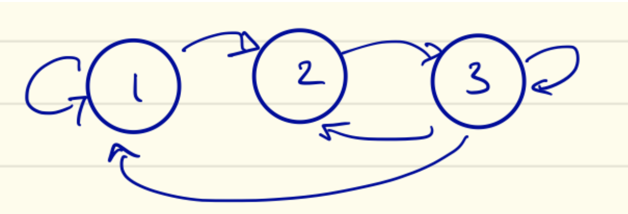
Definition: A chain that is not irreducible is **reducible**.

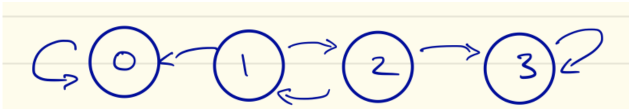
Your turn: Assume that each of the Markov chains given below have uniform transition probabilities. For each Markov chain

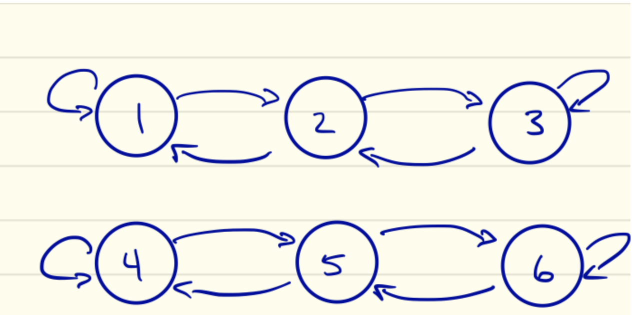
i. Classify the chain as reducible or irreducible

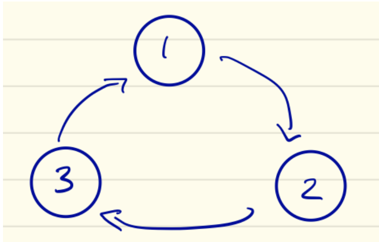
ii. Identify the transient states

iii. Identify the recurrent states









Long-run behavior

Definition. For irreducible, aperiodic Markov chains, the fraction of the time spent in each of the recurrent states is given by the **stationary distribution.** (a.k.a. steady state)

$\mathbf{s} = (s_1, s_2, \ldots, s_m)$ is a stationary distribution if

Key result: A Markov chain which starts out with a stationary distribution will stay in the stationary distribution forever.

Theorem. For any irreducible Markov chain:

1. A stationary distribution exists.
2. The stationary distribution is unique.
3. $s_i = 1/r_i$, where r_i is the expected number of steps required to return to state i, if starting at state i.
4. If Q^m is strictly positive (which implies aperiodic and recurrent) for some m, then

$$P(X_n = i) \to s_i \text{ as } n \to \infty$$