4.1 Introduction

Let $\{X_t,\, t \in T\}$ be a stochastic process.

- We say that X_t is the *state* of the process at *time* t.
- The *state space* is the set of all possible values that the random variables X_t can take on.

Definition 4.1.1. A *Markov chain* is a stochastic process $\{X_t,\, t = 0, 1, 2, \ldots\}$ with a countable state space that satisfies the so-called *Markov property*:

> Whenever the process is in state i, there is a fixed probability P_{ij} that it will next be in state j.

In other words:

$$P\{X_{t+1} = j \mid X_t = i, X_{t-1} = i_{t-1}, X_{t-2} = i_{t-2}, \ldots, X_1 = i_1, X_0 = i_0\}$$

$$= P\{X_{t+1} = j \mid X_t = i\}$$

for all $i_0, i_1, \ldots, i_{t-1}, i, j$, and all $t \geq 0$.

This means that, at any time t in a Markov chain, the distribution of the next state X_{t+1} is completely determined by the present state X_t; given the value of the present state X_t, the future state X_{t+1} is independent of the values of all of the past states $X_0, X_1, \ldots, X_{t-1}$.

The value

$$P_{ij} = P\{X_{t+1} = j \mid X_t = i\} \qquad (\text{for all } t)$$

is the probability that the process will, when in state i, next make a *transition* into state j; so the numbers P_{ij} are often called the *one-step transition probabilities*. It is often helpful to write them in a matrix:

e.g. with state space $\{1,2,3\}$

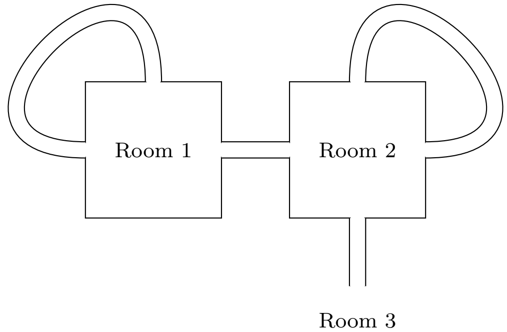
$$\begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} \end{array}$$

- Since probabilities are nonnegative, we have $P_{ij} \geq 0$
- If the process is in state i at a certain time t, then it must transition into *some* state j at time $t+1$, so

$$\sum_{\text{all } j} P_{ij} = 1 \qquad \text{for all } i.$$

row sums are 1.

Example 4.1.2. Suppose that Ernie is randomly placed in Room 1 or Room 2 of the maze below.



Ernie does not have a very good memory. Suppose that whenever Ernie is in Room 1 or Room 2, he randomly chooses one of the exits, and walks through the tunnel to the other side, independently of all of his other choices. Once he reaches Room 3, assume that he stays there indefinitely. Consider the process $\{X_t,\, t = 0, 1, 2, \ldots\}$, where X_t is the room that Ernie is in after walking through t tunnels.

(a) What is the state space of this process?

The state space is $\{1,2,3\}$.

(b) Explain why this process satisfies the Markov property.

Whenever Ernie is in Room i, there is a fixed probability that Ernie will next be in Room j.
(Regardless of how he got to Room i!)

(c) Write the transition probability matrix of the process.

state at time $t+1$

$$\begin{array}{ccc} & & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \text{state at time } t & \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

state 3 is an absorbing state.

Example 4.1.3. Suppose that the chance of sun tomorrow depends on previous weather conditions only through whether or not it is sunny today and not on past weather conditions. Suppose that

- if it is sunny today, then it will be sunny tomorrow with probability α; and
- if it is cloudy today, then it will be cloudy tomorrow with probability β.

Define a Markov chain that describes the weather conditions on day $t \geq 0$, and find its transition probability matrix.

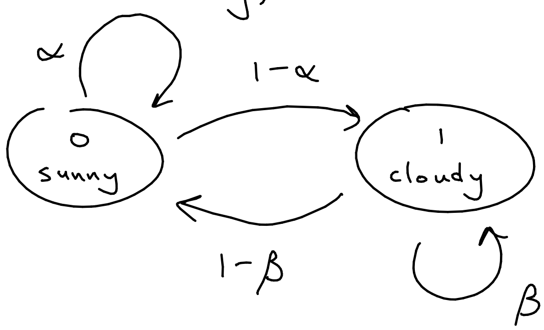
Say that

State 0: sunny

State 1: cloudy

Let X_t be the state on day t.

Then $\{X_t, t = 0, 1, 2, \ldots\}$ is a Markov Chain, since the weather tomorrow depends only on the weather today, and not on previous weather conditions.



Transition Probability Matrix:

$$\begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{bmatrix} \end{array}$$

Example 4.1.4. Suppose that the chance of sun tomorrow depends on previous weather conditions over the last *two* days (today and yesterday). Specifically, suppose that if it was...

- sunny for the past two days, then it will be sunny tomorrow with probability 0.9;
- sunny today but not yesterday, then it will be sunny tomorrow with probability 0.6;
- cloudy today but not yesterday, then it will be sunny tomorrow with probability 0.4;
- cloudy for the past two days, then it will be sunny tomorrow with probability 0.3.

(a) In this new setting, the two-state process defined in the previous example is no longer a Markov chain. Why?

The weather tomorrow depends not just on the weather today, but also on the weather yesterday.

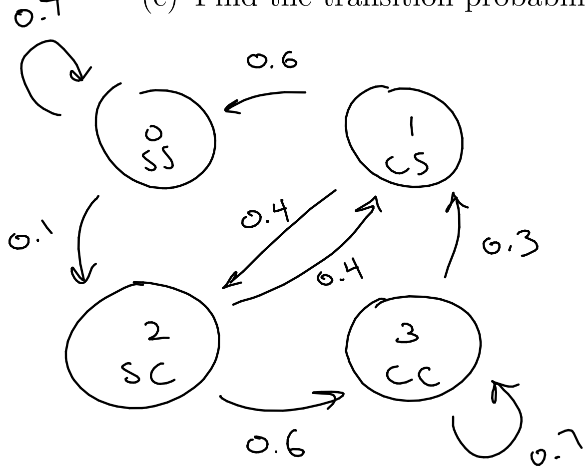
(b) Can we define a Markov chain that describes the weather conditions over time?

Let's say

State 0: It was sunny today and yesterday.
State 1: It was sunny today but not yesterday.
State 2: It was cloudy today but not yesterday.
State 3: It was cloudy today and yesterday.

Let X_t be the state at time t. Then $\{X_t, t=0,1,2,\ldots\}$ is a Markov chain, since whenever the process is in state i, there is a fixed probability that it will next be in state j.

(c) Find the transition probability matrix of this process.



$$\begin{array}{c c} & \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} 0.9 & 0 & 0.1 & 0 \\ 0.6 & 0 & 0.4 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.3 & 0 & 0.7 \end{bmatrix} \end{array}$$

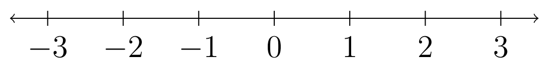
Example 4.1.5 (A Random Walk Model)**.** Consider a Markov chain with state space $\mathbb{Z}$, such that for all $i \in \mathbb{Z}$, we have

$$P_{ij} = \begin{cases} p_i, & \text{if } j = i+1; \\ q_i, & \text{if } j = i-1; \\ r_i, & \text{if } j = i; \\ 0, & \text{otherwise.} \end{cases}$$

$p_i + q_i + r_i = 1$

Imagine that the state of the process is the location of a (possibly intoxicated) person.

$X_t = 2$



$X_{t+1} = 1$ with prob. q_2

$X_{t+1} = 3$ with prob. p_2

$X_{t+1} = 2$ with prob. r_2

Then whenever the person is at location i, there are three possibilities:

- They move to the right with prob. p_i
- They move to the left with prob. q_i
- They stay where they are with prob. r_i

In the most simple random walks, p_i, q_i, and r_i are constants that do not depend on i.

Example 4.1.6 (A Gambling Model)**.** Consider a gambler who starts with some positive integer amount $\$X_0$, and at each play of the game, either wins \$1 with probability p or loses \$1 with probability $1-p$, independently of the outcomes of all other games. Suppose that our gambler quits playing when they go broke or they attain a fortune of $\$N$.

(a) Let X_t be the gambler's fortune after t plays. Explain why $\{X_t, t = 0, 1, 2, \ldots\}$ satisfies the Markov property.

Whenever the process is in state i, there is a fixed probability that the process is next in state j, regardless of how we got to state i. This is since all of the games are independent!

(b) Find the one-step transition probabilities, and write the transition probability matrix.

For $0 < i < N$,

$$P_{ij} = \begin{cases} p, & \text{if } j = i+1 \\ 1-p, & \text{if } j = i-1 \\ 0, & \text{otherwise} \end{cases}$$

$$P_{0j} = \begin{cases} 1, & \text{if } j = 0 \\ 0, & \text{otherwise} \end{cases}$$

$$P_{Nj} = \begin{cases} 1, & \text{if } j = N \\ 0, & \text{otherwise} \end{cases}$$

usually, we just write $P_{00} = P_{NN} = 1$

$$\begin{array}{c c} & \begin{array}{ccccccc} 0 & 1 & 2 & 3 & \cdots & N-1 & N \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ N-1 \\ N \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1-p & 0 & p & 0 & \cdots & 0 & 0 \\ 0 & 1-p & 0 & p & \ddots & & \\ \vdots & & \ddots & \ddots & \ddots & & \\ 0 & 0 & 0 & & 1-p & 0 & p \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \end{array}$$

4.2 The Chapman-Kolmogorov Equations

Note: When discussing a general Markov chain, we often let the state space be $\{0, 1, 2, \ldots\}$.

Recall that for a Markov chain, the one-step transition probabilities are defined by:

$$P_{ij} = P\{X_{t+1} = j \mid X_t = i\}$$

In words, P_{ij} is the probability that we will next be in state j given that we are presently in state i.

Extending this definition, we define the *n-step transition probabilities* as:

$$P_{ij}^{(n)} = P\{X_{t+n} = j \mid X_t = i\}$$

In words, $P_{ij}^{(n)}$ is the probability that, starting in state i, we will be in state j after n steps.

- Note that $P_{ij}^{(1)} = P_{ij}$

The *Chapman-Kolmogorov equations* provide a method for computing these n-step transition probabilities. They say that

$$P_{ij}^{(m+n)} = \sum_{\text{all } k} P_{ik}^{(m)} P_{kj}^{(n)}$$

here k is a state

Why does these equations hold?

- $P_{ik}^{(m)} P_{kj}^{(n)}$ represents: the probability that, starting in state i, we reach state j after $m+n$ steps through a path that visits state k after m steps.
- Since the process must go through some state after the mth transition, we sum these probabilities over all possible states k.

If we let $\mathbf{P^{(n)}}$ denote the matrix of n-step transition probabilities, then the Chapman-Kolmogorov equations tell us that

$$\mathbf{P}^{(m+n)} = P^{(m)} P^{(n)}$$

In particular, we have

$$\mathbf{P}^{(2)} = P^{(1+1)} = P^{(1)} P^{(1)} = P \cdot P = P^2$$

and by a straightforward proof by induction, we obtain

$$\mathbf{P}^{(n)} = P^{(n-1+1)} = P^{(n-1)} P^{(1)} = P^{n-1} \cdot P = P^n$$

Example 4.2.1. Consider the weather as a two-state Markov chain, as in Example 4.1.3. Suppose that

- if it is sunny today, then it will be sunny tomorrow with probability 0.8; and
- if it is cloudy today, then it will be cloudy tomorrow with probability 0.4.

(a) Write the one-step transition probability matrix.

sunny

$$P = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} \end{array}$$

cloudy

(b) Given that it is sunny today, find the probability that it will be sunny two days from now.

The two-step transition matrix is given by

$$P^{(2)} = P^2 = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix}$$

$$= \begin{bmatrix} 0.76 & 0.24 \\ 0.72 & 0.28 \end{bmatrix}$$

So the desired prob. is $P_{00}^{(2)} = 0.76$.

(c) Given that it is cloudy today, find the probability that it will be sunny four days from now.

next state

The four-step transition matrix is

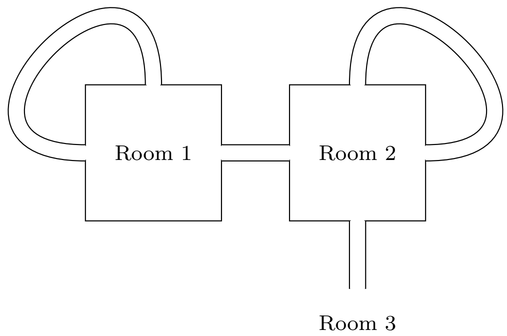
$$P^{(4)} = P^4 = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} 0.7504 & 0.2496 \\ 0.7488 & 0.2512 \end{bmatrix} \end{array}$$

present state

The desired probability is:

$$P_{10}^{(4)} = 0.7488$$

Example 4.2.2. Consider Ernie being placed in a maze again, as in Example 4.1.2.



Let X_t be the room that Ernie is in after walking through t tunnels.

(a) Given that Ernie is initially placed in Room 1, find the probability that he escapes by walking through *at most* 5 tunnels.

We want $P\{X_5 = 3 \mid X_0 = 1\} = P_{13}^{(5)}$

The one-step transition matrix is:
$$P = \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & & & \\ 2 & & \left[\begin{matrix} 2/3 & 1/3 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 1 \end{matrix}\right] & \\ 3 & & & \end{array}$$

So the five-step transition matrix is $P^5 = \begin{bmatrix} ? & ? & 0.35687 \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix}$

$\therefore$ The desired probability is 0.35687.

(b) Given that Ernie is initially placed in Room 1, find the probability that he escapes by walking through *exactly* 5 tunnels.

This probability can be computed as:

$$\underbrace{P_{13}^{(5)}}_{\text{prob. that he escapes in at most five steps}} - \underbrace{P_{13}^{(4)}}_{\text{prob that he escapes in at most four steps}} \approx 0.35687 - 0.27315 = 0.08372$$

Or as:

$$P_{12}^{(4)} \cdot P_{23} \approx (0.335) \cdot \frac{1}{4} = 0.08375$$

So far, we have only computed probabilities where the initial state is given or known. It could be the case that we only know the probability distribution of the initial state.

Let α_i denote the probability that $X_0 = i$. Then we can find the probability that $X_n = j$ by conditioning on X_0, as follows.

$$P\{X_n = j\} = \sum_{\text{all } i} P\{X_n = j \mid X_0 = i\} \cdot P\{X_0 = i\} = \sum_{\text{all } i} P_{ij}^{(n)} \alpha_i$$

Example 4.2.3. Find the probability that Ernie gets out of the maze in at most 5 steps if he is equally likely to be placed in either Room 1 or Room 2 at the beginning.

$$P\{X_5 = 3\} = P\{X_5 = 3 \mid X_0 = 1\} \cdot P\{X_0 = 1\} + P\{X_5 = 3 \mid X_0 = 2\} \cdot P\{X_0 = 2\}$$

$$= P_{13}^{(5)} \cdot \alpha_1 + P_{23}^{(5)} \cdot \alpha_2$$

$$\approx (0.35687)\tfrac{1}{2} + (0.58044) \cdot \tfrac{1}{2}$$

$$\approx 0.46865$$

Let $\vec{\alpha}$ denote the row vector whose ith entry is the probability $\alpha_i = P\{X_0 = i\}$:

$$\boldsymbol{\alpha} = [\alpha_1 \quad \alpha_2 \quad \alpha_3]$$

Then by the equations above, the product $\boldsymbol{\alpha}\mathbf{P}^n$ is a row vector whose jth entry is:

$$(\alpha P^n)_j = P\{X_n = j\}$$

Example 4.2.4. For $j \in \{1, 2, 3\}$, find the probability that Ernie is in Room j after 5 steps if he is equally likely to be placed in Room 1 or Room 2 at the beginning.

We have $\alpha = [\tfrac{1}{2}, \tfrac{1}{2}, 0]$,

so $\alpha P^5 = [0.2843, 0.2470, 0.4687]$

$P\{X_5 = 1\}$ $P\{X_5 = 2\}$ $P\{X_5 = 3\}$.

Example 4.2.5. Suppose that a fair coin is flipped repeatedly, and let N denote the number of flips until there is a run of three consecutive heads.

(a) Find the probability that N is at most 6.

Define a Markov chain with states 0, 1, 2, and 3, where $X_t = i$ if we have i consecutive heads after the ith flip, but once $X_t = 3$, we stay there indefinitely.

Then the transition probability matrix is

$$\begin{array}{c|cccc} & 0 & 1 & 2 & 3 \\ \hline 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 2 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 3 & 0 & 0 & 0 & 1 \end{array}$$

We want
$$P\{N \le 6\} = P\{X_6 = 3 \mid X_0 = 0\}$$
$$= P_{03}^{(6)}$$
$$= 0.3125 \quad (\text{using Sage})$$

(b) Suppose instead that the coin comes up heads with probability p, and tails with probability $q = 1 - p$. Find the probability that N is at most 6.

Now, the transition probability matrix is

$$P = \begin{bmatrix} 1-p & p & 0 & 0 \\ 1-p & 0 & p & 0 \\ 1-p & 0 & 0 & p \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Writing $q = 1-p$, we have

$$P_{03}^{(6)} = (pq + q^2)p^3 + p^3 q + (p^3 q + p^2 q^2 + (pq + q^2)pq)p^2$$
$$+ p^3$$

4.3 Classification of States

The definitions and statements in this section are made for all Markov chains, but we omit this quantifier for ease of reading.

Definition 4.3.1. We say that state j is *accessible* from state i if

$$P_{ij}^{(n)} > 0 \quad \text{for some} \quad n \geq 0.$$

Notice that if state j is *not* accessible from state i, then $P_{ij}^{(n)} = 0$ for all $n \geq 0$.

So $P\{\text{ever be in state } j \mid \text{start in } i\} = P\{\bigcup_{n=0}^{\infty} X_n = j \mid X_0 = i\} \leq \sum_{n=0}^{\infty} P\{X_n = j \mid X_0 = i\} = \sum_{n=0}^{\infty} P_{ij}^{(n)} = 0$

Thus, if j is not accessible from i, and the process reaches state i at some point, then it will *never again* enter state j.

Definition 4.3.2. We say that state i *communicates* with state j, and write $i \leftrightarrow j$, if

j is accessible from i and i is accessible from j.

Observation 4.3.3. The relation of communication is an equivalence relation. That is, the relation $\leftrightarrow$ is

(a) Reflexive: $\forall$ states i, $i \leftrightarrow i$

(b) Symmetric: $\forall$ states i, j, if $i \leftrightarrow j$, then $j \leftrightarrow i$

(c) Transitive: $\forall$ states i, j, k, if $i \leftrightarrow j$ and $j \leftrightarrow k$, then $i \leftrightarrow k$.

Proof of (c). Let i, j, k be states such that $i \leftrightarrow j$ and $j \leftrightarrow k$.

Since j is accessible from i, there is some number n such that $P_{ij}^{(n)} > 0$. Since k is accessible from j, there is some number m such that $P_{jk}^{(m)} > 0$. $\therefore$ By the Chapman-Kolmogorov equations, we have $P_{ik}^{(n+m)} = \sum_{\text{all } r} P_{ir}^{(n)} P_{rk}^{(m)} \geq P_{ij}^{(n)} P_{jk}^{(m)} > 0$

$\therefore$ k is accessible from i.

By a similar argument, i is accessible from k. $\therefore$ $i \leftrightarrow k$. $\square$

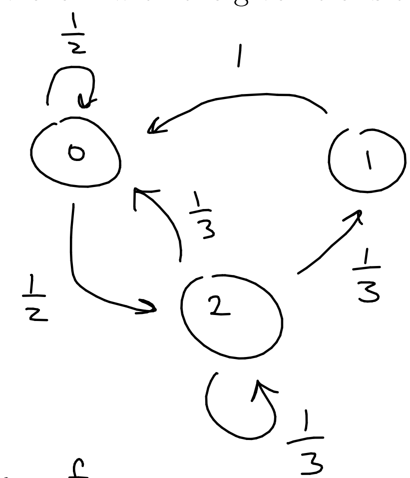
Thus, the relation of communication partitions the state space into disjoint sets called *equivalence classes*, or just *classes* for short.

- Any two states that communicate belong to the same class
- Any two states that do not communicate belong to different classes.

A Markov chain with only one class is said to be irreducible.

Example 4.3.4. Find the classes of the Markov chain with the given transition probability matrix.

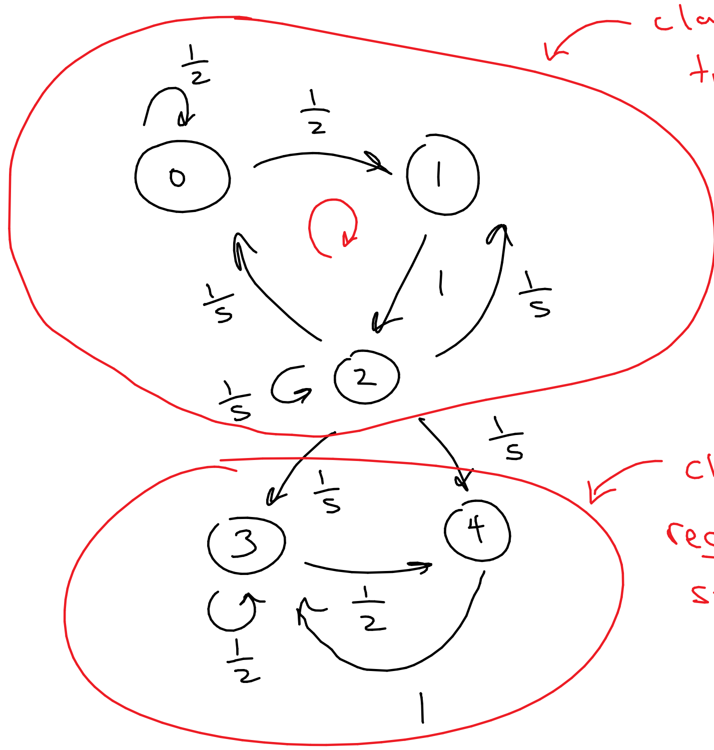
(a)
$$\begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array}\begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \end{array}$$



We check that all pairs of states communicate, so this Markov chain has just one class.

It is irreducible.

(b)
$$\begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$



class of transient states

class of recurrent states

The classes are $\{0,1,2\}$ and $\{3,4\}$.

all pairs here can communicate

all pairs here can communicate

but no way to get from $\{3,4\}$ to $\{0,1,2\}$.

Definition 4.3.5. For any state i, we let f_i denote the probability that, starting in state i, the process ever reenters state i. That is

there exists.

$$f_i = P\{\text{even return to } i \mid \text{start in } i\} = P\{\exists\, n>0 \text{ such that } X_n = i \mid X_0 = i\}$$

State i is said to be

English meaning: occurring often or repeatedly.

- recurrent if $f_i = 1$
- transient if $f_i < 1$

In English, the word "transient" means lasting only for a short time. ☺

Observation 4.3.6. If state i is recurrent then, starting in state i, the number of times that the process reenters state i is infinite.

Proof. Suppose that state i is recurrent and that we start in state i. Then with probability 1, when we leave state i, we will eventually return to state i.

By the Markov property, when we return to state i, we are right back where we started, and we will return to state i again with probability 1.

So we leave and return again and again, ad infinitum. $\square$

Observation 4.3.7. If state i is transient then, starting in state i, the number of times that the process reenters state i is finite.

In fact, the total number of times that we visit state i is a geometric r.v. with mean $\frac{1}{1-f_i}$.

Proof. Suppose that state i is transient. Then each time we enter state i, when we leave, the probability that we return is $f_i < 1$. $\therefore$ Every time we leave state i, there is some positive probability that we never come back ☹, namely $1-f_i$.

Therefore, starting in state i, the probability that we will visit state i exactly n times is

$$f_i^{\,n-1}(1-f_i).$$

$\therefore$ The number of times that we visit state i is geometric with parameter $(1-f_i)$, so has mean $1/(1-f_i)$. $\square$

Corollary 4.3.8. *Every finite-state Markov chain has a* recurrent state.

Proof. Suppose otherwise that there is a finite-state Markov chain with only transient states. Then every state is visited only finitely many times. But this is impossible, since the process is infinite; it goes on forever. □

Proposition 4.3.9. *State i is recurrent if and only if* $\sum_{n=1}^{\infty} P_{ii}^{(n)}$ is infinite.

Proof. Suppose that state i is recurrent.

Then by Obervation 4.3.6, the number of times that we return to state i, given that we start in state i, is infinite.

Let
$$I_n = \begin{cases} 1, & \text{if } X_n = i \\ 0, & \text{if } X_n \neq i, \end{cases}$$

so that $\sum_{n=1}^{\infty} I_n$ is the number of periods that the process is in state i. Then
$$E\Big[\sum_{n=1}^{\infty} I_n \,\Big|\, X_0 = i\Big] = \sum_{n=1}^{\infty} E[I_n \mid X_0 = i]$$
$$= \sum_{n=1}^{\infty} P\{X_n = i \mid X_0 = i\}$$
$$= \sum_{n=1}^{\infty} P_{ii}^{(n)}$$

expected number of times that we return to state i, which is infinite!

$\therefore$ The sum $\sum_{n=1}^{\infty} P_{ii}^{(n)} = \infty$. □

Corollary 4.3.10. *State i is transient if and only if* $\sum_{n=1}^{\infty} P_{ii}^{(n)}$ is finite.

Proof. Suppose that state i is transient.

By an argument similar to the one used above, we have
$$E\Big[\sum_{n=1}^{\infty} I_n \,\Big|\, X_0 = i\Big] = \sum_{n=1}^{\infty} P_{ii}^{(n)}$$

But from Observation 4.3.7, we know that
$$E\Big[\sum_{n=1}^{\infty} I_n \,\Big|\, X_0 = i\Big] = \frac{1}{1-f_i}.$$

$\therefore$ $\sum_{n=1}^{\infty} P_{ii}^{(n)} = \dfrac{1}{1-f_i}$, which is finite. □

Corollary 4.3.11. *If state i is recurrent and $i \leftrightarrow j$, then state j is* recurrent.

In other words: If i is recurrent, then so are all states in its class.

Proof. Suppose that i is recurrent, and that $i \leftrightarrow j$.

Since $i \leftrightarrow j$, there is some number m_1 such that $p_{ji}^{(m_1)} > 0$ and some number m_2 such that $p_{ij}^{(m_2)} > 0$.

$\therefore$ For all $m_3 > 0$, we have

$$p_{jj}^{(m_1+m_2+m_3)} \geq p_{ji}^{(m_1)} p_{ii}^{(m_3)} p_{ij}^{(m_2)}$$

$\therefore$ We have

$$\sum_{n=1}^{\infty} p_{jj}^{(n)} \geq \sum_{m_3=1}^{\infty} p_{ji}^{(m_1)} p_{ii}^{(m_3)} p_{ij}^{(m_2)}$$

$$= p_{ji}^{(m_1)} p_{ij}^{(m_2)} \cdot \sum_{m_3=1}^{\infty} p_{ii}^{(m_3)}$$

Since $\sum_{m_3=1}^{\infty} p_{ii}^{(m_3)}$ is infinite, as i is recurrent, we conclude that $\sum_{n=1}^{\infty} p_{jj}^{(n)}$ is also infinite.

$\therefore$ State j is also recurrent. $\square$

Corollary 4.3.12. *If state i is transient, and $i \leftrightarrow j$, then state j is* transient

In other words: If state i is transient, then so are all states in its class.

Proof. Suppose that state i is transient, and that $i \leftrightarrow j$.

Suppose towards a contradiction that j is recurrent.

Then by Corollary 4.3.11, since $j \leftrightarrow i$, state i must also be recurrent. But this contradicts the assumption that state i is transient. $\square$

Let's summarize what we've proven about recurrent and transient states.

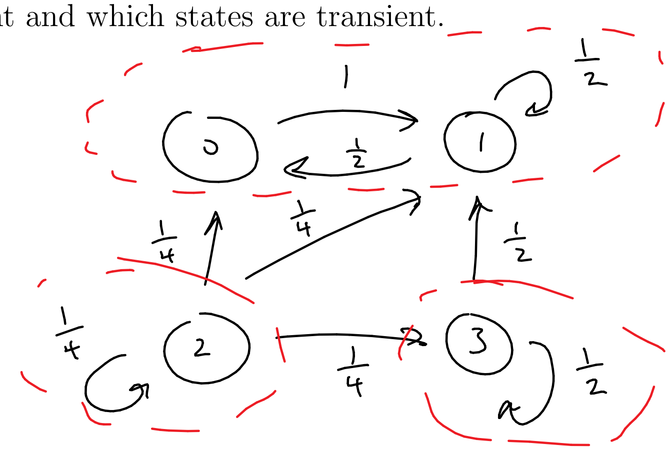
- State i is recurrent if and only if
  - $f_i = 1$
  - starting in state i, we return to state i infinitely many times.
  - $\sum_{n=1}^{\infty} p_{ii}^{(n)}$ is infinite
- State i is transient if and only if
  - $f_i < 1$
  - starting in state i, we return to state i only finitely many times.
  - $\sum_{n=1}^{\infty} p_{ii}^{(n)}$ is finite
- If i and j communicate, then they are either both recurrent or both transient.
  - In other words, recurrence and transience are *class properties.*
  - If one state in a class is recurrent, then all states in that class are recurrent.
  - If one state in a class is transient, then " " " " " " transient.
- Every finite-state Markov chain has a recurrent state.

Corollary 4.3.13. *Every finite-state irreducible Markov chain has* all recurrent states.

a single class!
i.e., all pairs of states communicate.

Example 4.3.14. Consider the Markov chain with the given transition probability matrix. Determine which states are recurrent and which states are transient.

(a)
$$\begin{array}{c c} & \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \end{array}$$

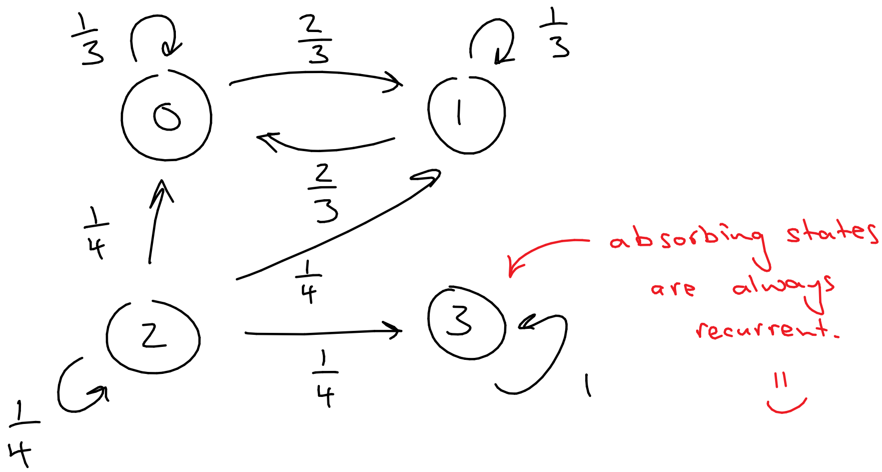


The classes are $\{0,1\}$, $\{2\}$, and $\{3\}$.

All states in the class $\{0,1\}$ are recurrent.

States 2 and 3 are transient, since every time we leave these states, there is a nonzero probability that we will never come back.

(b)
$$\begin{array}{c|cccc} & 0 & 1 & 2 & 3 \\ \hline 0 & \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ 1 & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 2 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 3 & 0 & 0 & 0 & 1 \end{array}$$



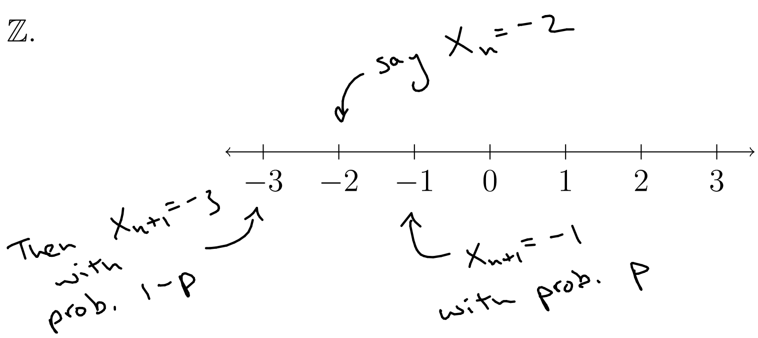
The classes are $\{0,1\}$, $\{2\}$, and $\{3\}$.

The states 0 and 1 are recurrent, the state 2 is transient, and the state 3 is recurrent.

Example 4.3.15 (A random walk)**.** Consider a Markov chain whose state space is $\mathbb{Z}$, and whose transition probabilities are given by

$$P_{i,i+1} = p \quad \text{and} \quad P_{i,i-1} = 1 - p, \quad \text{where } 0<p<1,$$

for all $i \in \mathbb{Z}$.



(a) Determine the classes of the state space.

All pairs of states communicate!

For any states i and j with $i<j$, we have $P_{ij}^{(j-i)} \geq p^{j-i} > 0$

and $P_{ji}^{(j-i)} \geq (1-p)^{j-i} > 0$.

$\therefore$ There is only one class, i.e. the chain is <u>irreducible</u>.

(b) Are the states of this Markov chain recurrent or transient?

Since all states are in the same class, we just need to determine if a single state, say 0, is recurrent or transient.

We know that state 0 is recurrent if $\sum_{n=1}^{\infty} P_{00}^{(n)} = \infty$, and transient if $\sum_{n=1}^{\infty} P_{00}^{(n)}$ is finite.

So let's try to determine $\sum_{n=1}^{\infty} P_{00}^{(n)}$.

Note that $P_{00}^{(2n-1)} = 0$, since it takes an <u>even</u> number of steps, starting from 0, to return to 0.

If we return to 0 (starting from 0) in $2n$ steps, then we must have taken n steps to the right and n steps to the left. So

$$P_{00}^{(2n)} = \binom{2n}{n} p^n (1-p)^n.$$

So

$$\sum_{n=1}^{\infty} P_{00}^{(n)} = \sum_{n=1}^{\infty} P_{00}^{(2n)} = \sum_{n=1}^{\infty} \binom{2n}{n} p^n (1-p)^n.$$

<u>Stirling's Approximation</u>: $n! \sim n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi}$, where $a_n \sim b_n$ means $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$

So we obtain

$$\binom{2n}{n} = \frac{(2n)!}{n!\,n!} \sim \frac{(2n)^{2n+\frac{1}{2}} \cancel{e^{-2n}} \sqrt{2\pi}}{n^{2n+1} \cancel{e^{-2n}} \cdot 2\pi} = \frac{2^{2n}}{\sqrt{n\pi}}$$

$$\therefore \binom{2n}{n} p^n (1-p)^n \sim \frac{2^{2n}}{\sqrt{n\pi}} p^n (1-p)^n = \frac{[4p(1-p)]^n}{\sqrt{n\pi}}$$

By the Limit Comparison Test, the series $\sum_{n=1}^{\infty} P_{00}^{(n)}$ converges if and only if $\sum_{n=1}^{\infty} \frac{[4p(1-p)]^n}{\sqrt{n\pi}}$ converges.

If $p = \frac{1}{2}$, then

$$\sum_{n=1}^{\infty} \frac{[4p(1-p)]^n}{\sqrt{n\pi}} = \sum_{n=1}^{\infty} \frac{1}{\sqrt{n\pi}} = \frac{1}{\sqrt{\pi}} \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}},$$

Symmetric Random Walk.

which is divergent. So when $p = \frac{1}{2}$, all states are <u>recurrent</u>.

If $p \neq \frac{1}{2}$, then $4p(1-p) < 1$ (use differential calculus to show that $4p(1-p)$ is maximized at $p = \frac{1}{2}$)

Then $\sum_{n=1}^{\infty} \frac{[4p(1-p)]^n}{\sqrt{n\pi}}$ converges by comparison with the geometric series $\sum_{n=1}^{\infty} [4p(1-p)]^n$. So when $p \neq \frac{1}{2}$, all states are <u>transient</u>.

4.4 Long-Run Proportions and Limiting Probabilities

For states $i \neq j$, let $f_{i,j}$ denote the probability that, starting in state i, the Markov chain will ever enter state j. That is,

$$f_{i,j} = P\{\exists\, n>0 \text{ such that } X_n = j \mid X_0 = i\}.$$

Proposition 4.4.1. *If state i is recurrent and $i \leftrightarrow j$, then $f_{i,j} = 1$.*

Sketch of Proof. Since $i \leftrightarrow j$, there is some number n such that $P_{ij}^{(n)} > 0$. Since i is recurrent, starting in state i, we visit state i infinitely many times. Each time we visit state i, there is a positive probability that we visit state j exactly n steps later. Therefore, we must eventually visit state j. □

Suppose that state j is recurrent.

- Then by definition, starting in state j, the probability that the process returns to state j is: $f_j = 1$
- Let
$$m_j = E[N_j \mid X_0 = j], \quad \text{where } N_j = \min\{n > 0\colon\ X_n = j\}.$$
- In other words, m_j is the expected number of steps, starting from state j, until we return to state j for the first time.

Definition 4.4.2. We say that a recurrent state j is

- *positive recurrent* if m_j is finite
- *null recurrent* if m_j is infinite.

Now suppose that the Markov chain is irreducible and recurrent, i.e., has all recurrent states.

- In the long-run, what proportion of time does the process spend in state j?
- Define $\pi_j = \lim\limits_{n\to\infty} \dfrac{n}{\text{step on which state } j \text{ is visited for the } n\text{th time}}$.
- We want to find this long-run proportion π_j.

Proposition 4.4.3. *In an irreducible and recurrent Markov chain, regardless of the initial state, we have*

$$\pi_j = \frac{1}{m_j}$$

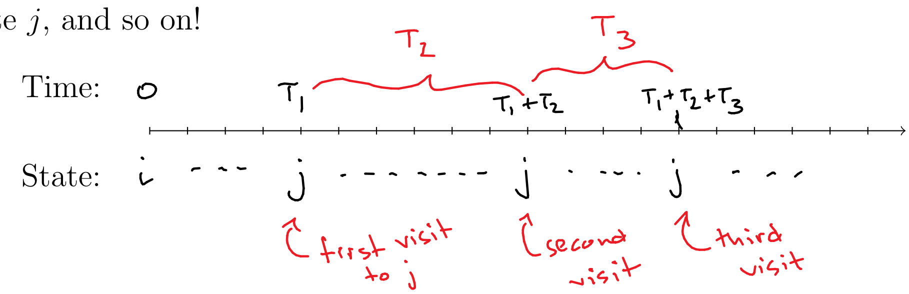
for all states j.

For the proof, recall the **Strong Law of Large Numbers**: If $X_1, X_2, \ldots$ are a sequence of independent and identically distributed random variables with common mean μ, then with probability 1, we have

$$\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu$$

Proof. Suppose that the Markov chain starts in state i.

- Let T_1 be the number of transitions until the process first enters state j.
- Let T_2 be the number of additional transitions (after T_1) until the process returns to state j.
- Let T_3 be the number of additional transitions (after $T_1 + T_2$) until the process returns to state j, and so on!



Note that:

- T_1 is finite by Proposition 4.4.1.
- By the Markov property, the random variables $T_2, T_3, \ldots$ are independent and identically distributed with mean: m_j
- The nth visit to state j occurs at time: $T_1 + T_2 + T_3 + \cdots + T_n$

It follows that π_j, the long-run proportion of time that the chain is in state j, is

$$\pi_j = \lim_{n\to\infty} \frac{n}{T_1 + T_2 + \cdots + T_n} = \lim_{n\to\infty} \frac{1}{\frac{T_1 + T_2 + \cdots + T_n}{n}}$$

$$= \lim_{n\to\infty} \frac{1}{\frac{T_1}{n} + \frac{T_2 + T_3 + \cdots + T_n}{n}}$$

$$= \frac{1}{\lim_{n\to\infty} \frac{T_1}{n} + \lim_{n\to\infty} \frac{T_2 + T_3 + \cdots + T_n}{n-1} \cdot \frac{n-1}{n}}$$

Since T_1 is finite

$$= \frac{1}{0 + m_j \cdot 1}$$

by the Strong Law of Large Numbers!

$$= \frac{1}{m_j}$$

$\square$

Corollary 4.4.4. *In an irreducible and recurrent Markov chain, state j is positive recurrent if and only if $\pi_j > 0$.*

Proof. By definition, state j is positive recurrent if and only if m_j is finite, which by Proposition 4.4.3, occurs if and only if $\pi_j = \frac{1}{m_j} > 0$. □

Proposition 4.4.5. *In an irreducible and recurrent Markov chain, either all states are positive recurrent, or all states are null recurrent.*

Proof. Suppose that state i is positive recurrent in an irreducible and recurrent Markov chain.
Let j be any other state. We want to show that state j is also positive recurrent. Since state i is positive recurrent, we have $\pi_i > 0$ by Corollary 4.4.4.
Since $i \longleftrightarrow j$, there is some number n such that $P_{ij}^{(n)} > 0$.
But then $\pi_j \geq \pi_i \cdot P_{ij}^{(n)} > 0$

long-run proportion of time in state i

long-run proportion of time that we visit state j n steps after visiting state i.

$\therefore$ State j is positive recurrent by Corollary 4.4.4. □

Corollary 4.4.6. *Every irreducible finite-state Markov chain is positive recurrent, i.e., has all positive recurrent states.*

Proof. We know that every irreducible finite-state Markov chain is recurrent, and by Proposition 4.4.5, all states must be positive recurrent, or all states are null recurrent. Since there are only finitely states, some state must have a positive long-run proportion, hence <u>all</u> states must be positive recurrent. □

To determine the long-run proportions for all states, note that the product

$$\underbrace{\pi_i}_{\text{long-run proportion of time in state } i} \underbrace{P_{i,j}}_{\text{long-run proportion of transitions from state } i \text{ that go to state } j}$$

is the long-run proportion of transitions from state i to state j. Summing this quantity over all states i gives the long-run proportion of time spent in state j:

$$\pi_j = \sum_{\text{all } i} \pi_i P_{i,j}$$

Indeed, one can prove the following important result.

Theorem 4.4.7. *Consider an irreducible Markov chain. If the chain is positive recurrent, then the long-run proportions are the unique solution of the equations*

$$\begin{cases} \pi_j = \displaystyle\sum_{\text{all } i} \pi_i P_{i,j} \\ \displaystyle\sum_{\text{all } j} \pi_j = 1 \end{cases}$$

Further, if there is no solution of the above equations, then the Markov chain is either transient or null recurrent, and $\pi_j = 0$ for all j.

Example 4.4.8. Find the long-run proportions for the Markov chain with transition probability matrix

$$[\pi_0 \quad \pi_1] \quad P = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \end{array}$$

This chain is positive recurrent, so there must be some solution to the equations of Theorem 4.4.7.

So we solve:

$$\begin{cases} \pi_0 = 0.8\pi_0 + 0.3\pi_1 & (1) \\ \pi_1 = 0.2\pi_0 + 0.7\pi_1 & (2) \\ \pi_0 + \pi_1 = 1 & (3) \end{cases}$$

From (1), we get $0.2\pi_0 = 0.3\pi_1$

$$\Rightarrow \pi_0 = 1.5\pi_1$$

Substituting into (3), we get

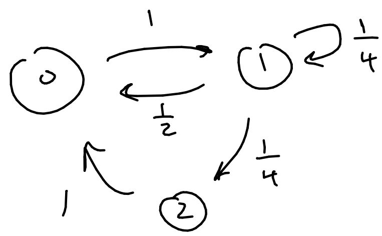
$$1.5\pi_1 + \pi_1 = 1$$

$$\Rightarrow \pi_1 = \frac{2}{5}$$

Back-substituting, we find $\pi_0 = \frac{3}{5}$.

Example 4.4.9. Find the long-run proportions for the Markov chain with transition probability matrix

$$P = \begin{array}{c c} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 1 & 0 & 0 \end{bmatrix} \end{array}$$



Note that all states communicate, and since there are only finitely many states, they must all be positive recurrent. $\therefore$ There is a unique solution to the equations:

$$\begin{cases} \pi_0 = 0\cdot\pi_0 + \frac{1}{2}\cdot\pi_1 + 1\cdot\pi_2 \\ \pi_1 = \pi_0 + \frac{1}{4}\pi_1 \\ \pi_2 = \frac{1}{4}\pi_1 \\ \pi_0+\pi_1+\pi_2 = 1 \end{cases}$$

We use Sage to find

$$\pi_0 = \frac{3}{8}, \quad \pi_1 = \frac{1}{2}, \quad \pi_2 = \frac{1}{8}.$$

The long-run proportions π_j are often called *stationary probabilities*. For if we start in state j with probability π_j, then the probability that we are in state j after n steps is still π_j.

Proposition 4.4.10 (Long-run proportions are stationary probabilities)**.** *In an irreducible, positive recurrent Markov chain, if*

$$P\{X_0 = j\} = \pi_j \quad \textit{for all } j,$$

then

$$P\{X_n = j\} = \pi_j \quad \text{for all } j, \text{ all } n \geq 0.$$

Proof. Suppose that $P\{X_0 = j\} = \pi_j$ for all j. We want to show that for any $n \geq 0$, we have $P\{X_n = j\} = \pi_j$ for all j. We proceed by induction on n.

- Base Case: The conclusion is immediate when $n = 0$.
- Inductive Hypothesis: Suppose for some $n > 0$ that
$$P\{X_{n-1} = j\} = \pi_j \quad \text{for all } j.$$
- Inductive Step: Then
$$\begin{aligned} P\{X_n = j\} &= \sum_{\text{all } i} P\{X_n = j \mid X_{n-1} = i\}\cdot P\{X_{n-1} = i\} \\ &= \sum_{\text{all } i} P_{ij}\,\pi_i \\ &= \sum_{\text{all } i} \pi_i P_{ij} \\ &= \pi_j, \quad \text{for all } j. \end{aligned}$$

by conditioning on the previous state

by Theorem 4.4.7

$\therefore$ By math induction, if $P\{X_0 = j\} = \pi_j$ for all j, then $P\{X_n = j\} = \pi_j$ for all j, all $n \geq 0$.

Recall that the Markov chain with transition probability matrix

$$P = \begin{array}{c c} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 1 & 0 & 0 \end{bmatrix} \end{array}$$

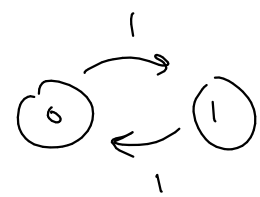
had long-run proportions

$$[\pi_0, \pi_1, \pi_2] = \left[\tfrac{3}{8}, \tfrac{1}{2}, \tfrac{1}{8}\right].$$

What happens when we take large powers of P?

For any initial state i and any state j, it appears that

$$\lim_{n\to\infty} P_{ij}^{(n)} = \pi_j$$



Does this always occur? Consider the Markov chain with transition probability matrix

$\pi_0 = 0 \cdot \pi_0 + \pi_1$

$\pi_1 = 1 \cdot \pi_0 + 0 \cdot \pi_1$

$\pi_0 + \pi_1 = 1$

$$P = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{array}$$

$$P^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- The limiting proportions are: $\pi_0 = \pi_1 = \frac{1}{2}$
- But

  $P^n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ if n is even, while $P^n = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ if n is odd.

However, this Markov chain is *periodic*.

- We say that state i is *periodic* if there is some number $d > 1$ such that

  $P_{ii}^{(n)} > 0$ only if n is a multiple of d.

- A Markov chain with no periodic states is called aperiodic.
- An irreducible, positive recurrent, aperiodic Markov chain is called ergodic.

Theorem 4.4.11 (Long-run proportions are limiting probabilities). *For all states j in an irreducible, positive recurrent, aperiodic Markov chain, we have*

$$\lim_{n\to\infty} P\{X_n = j\} = \pi_j$$

regardless of the initial state i.

Example 4.4.12. Lucas and Justin are playing a sequence of rallies in squash, which begin when one of them serves. Suppose that Lucas wins each rally that he serves with probability p, and wins each rally that Justin serves with probability q. The winner of a rally becomes the server of the next rally.

(a) Find the long-run proportion of time that Lucas serves.

(b) Find the long-run proportion of rallies that Lucas wins.

(c) Find the long-run proportion of rallies on which the serve changes.

Consider the Markov chain where state 0 represents Lucas serving and state 1 represents Justin serving. Then the transition matrix is

$$P = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} p & 1-p \\ q & 1-q \end{bmatrix} \end{array}$$

(a) To find the long-run proportion π_0, we solve

$$\pi_0 = p\pi_0 + q\pi_1 \quad (1)$$

$$\pi_1 = (1-p)\pi_0 + (1-q)\pi_1 \quad \text{(don't really need this one)}$$

$$\pi_0 + \pi_1 = 1 \quad (2)$$

From (1), we find $\pi_0 = \frac{q}{1-p}\pi_1$

Subbing into (2), we find $\frac{q}{1-p}\pi_1 + \pi_1 = 1$

$$\Rightarrow \left(\frac{q+1-p}{1-p}\right)\pi_1 = 1$$

$$\Rightarrow \pi_1 = \frac{1-p}{q+1-p}$$

$$\therefore \pi_0 = \frac{q}{q+1-p}$$

$\therefore$ The long-run proportion of time that Lucas serves is $\pi_0 = \frac{q}{q+1-p}$.

Since Lucas serves the next rally every time he wins, the long-run proportion of rallies that he wins is also π_0.

Finally, the long-run proportion of rallies on which the serve changes is $\underbrace{\pi_0 \cdot (1-p)}_{\text{Lucas serves and loses}} + \underbrace{\pi_1 \cdot q}_{\text{Justin serves and loses}}$ $\square$.

Recall the following:

- A recurrent state is...
  - *positive recurrent* if and only if m_i is finite, or equivalently, $\pi_i > 0$.
  - *null recurrent* if and only if m_i is infinite, or equivalently, $\pi_i = 0$.
- In an irreducible and recurrent Markov chain, either

  all states are positive recurrent or all states are null recurrent.
- The *symmetric random walk* has state space $\mathbb{Z}$ and transition probabilities

$$P_{i,i+1} = \tfrac{1}{2} = P_{i,i-1} \quad \text{for all } i \in \mathbb{Z}.$$

Example 4.4.13. Show that every state in the symmetric random walk is null recurrent.

It suffices to show that there are no solutions to the system

$$\begin{cases} \pi_j = \sum_{\text{all } i} \pi_i P_{ij} \\ \sum_{\text{all } i} \pi_i = 1 \end{cases}$$

The long-run proportions do not depend on the starting state, so suppose that we start in state 0.

Then by symmetry, $\pi_1 = \pi_{-1}$, $\pi_2 = \pi_{-2}$, and so on.

Suppose that $\pi_j = \sum_{\text{all } i} \pi_i P_{i,j}$ for all j.

That is,

$$\begin{cases} \vdots \\ \pi_{-1} = \frac{1}{2}\pi_{-2} + \frac{1}{2}\pi_0 \\ \pi_0 = \frac{1}{2}\pi_{-1} + \frac{1}{2}\pi_1 \\ \pi_1 = \frac{1}{2}\pi_0 + \frac{1}{2}\pi_2 \\ \vdots \end{cases}$$

Say $\pi_1 = \pi_{-1} = c$. Then $\pi_0 = \frac{1}{2}\pi_{-1} + \frac{1}{2}\pi_1 = \frac{1}{2}c + \frac{1}{2}c = c$

So $\pi_0 = \pi_1 = \pi_{-1} = c$. Then $\pi_1 = \frac{1}{2}\pi_0 + \frac{1}{2}\pi_2$

$$\Rightarrow c = \tfrac{1}{2}c + \tfrac{1}{2}\pi_2$$

$$\Rightarrow \pi_2 = c.$$

Similarly, $\pi_{-2} = c$, $\pi_3 = c$, $\pi_{-3} = c$, ...

But now $\sum_{\text{all } i} \pi_i \neq 1$ (if $c > 0$, it diverges, and if $c = 0$, then it is 0.)

$\therefore$ The system above has no solutions, and by Thm 4.4.7, all states are null recurrent and have long-run prop. 0.

4.5 The Gambler's Ruin Problem

Example 4.5.1. Consider a gambler who at each play of the game has probability p of winning one unit and probability $q = 1 - p$ of losing one unit, independently of all other plays. The gambler plays the game repeatedly until they either reach a fortune of N units, or they lose it all. Let X_n denote the gambler's fortune at time n.

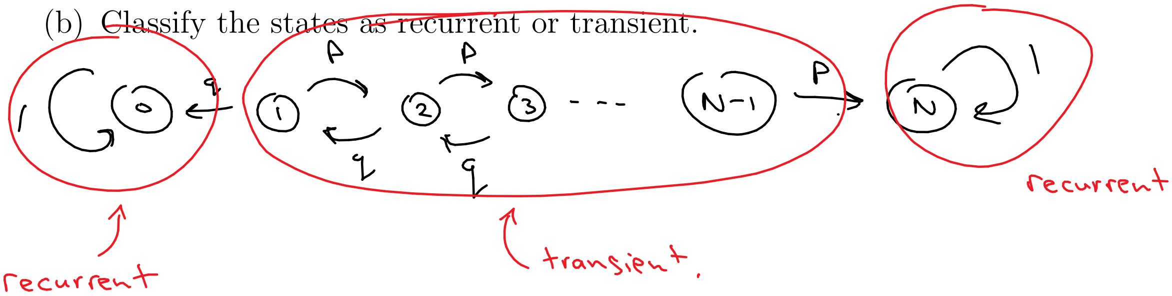
in $(0,1)$

(a) What are the transition probabilities?

The state space is $\{0, 1, 2, \dots, N\}$ and the transition probs are: $P_{00} = 1 = P_{NN}$

and $P_{i,i+1} = p$, $P_{i,i-1} = q = 1-p$ for all $i \in \{1, 2, \dots, N-1\}$.

(b) Classify the states as recurrent or transient.



There are three classes: $\{0\}$, $\{1, 2, \dots, N-1\}$, and $\{N\}$.

The states 0 and N are recurrent, while the class $\{1, 2, \dots, N-1\}$ is transient.

(c) Let P_i, $i = 0, 1, 2, \dots, N$, denote the probability that, starting with i units, the gambler will reach a fortune of N (before they lose it all). Find P_i.

Suppose that we start with i units. Conditioning on the outcome of the first play, we find

$$P_i = p \cdot P_{i+1} + q \cdot P_{i-1} \quad \text{for } i = 1, 2, \dots, N-1$$

since $p + q = 1$

$$\Longrightarrow \quad pP_i + qP_i = p \cdot P_{i+1} + q \cdot P_{i-1}$$

$$\Longrightarrow \quad p \cdot P_{i+1} - p \cdot P_i = q \cdot P_i - q \cdot P_{i-1}$$

$$\Longrightarrow \quad P_{i+1} - P_i = \frac{q}{p}\left(P_i - P_{i-1}\right)$$

We also have $P_0 = 0$ and $P_N = 1$.

Since $P_0 = 0$, we have

$$P_2 - P_1 = \frac{q}{p}(P_1 - P_0) = \frac{q}{p}P_1$$

$$P_3 - P_2 = \frac{q}{p}(P_2 - P_1) = \left(\frac{q}{p}\right)^2 P_1$$

$$P_4 - P_3 = \frac{q}{p}(P_3 - P_2) = \left(\frac{q}{p}\right)^3 P_1$$

$$\vdots$$

$$P_i - P_{i-1} = \frac{q}{p}(P_{i-1} - P_{i-2}) = \left(\frac{q}{p}\right)^{i-1} P_1$$

$$\vdots$$

$$P_N - P_{N-1} = \left(\frac{q}{p}\right)^{N-1} P_1$$

Adding the first $i-1$ of these equations, we obtain

$$P_i - P_1 = P_1\left[\frac{q}{p} + \left(\frac{q}{p}\right)^2 + \left(\frac{q}{p}\right)^3 + \dots + \left(\frac{q}{p}\right)^{i-1}\right]$$

$$\Rightarrow \quad P_i = P_1\left[1 + r + r^2 + r^3 + \dots + r^{i-1}\right] \quad \text{where } r = \frac{q}{p}.$$

$$\therefore P_i = \begin{cases} P_1\left(\frac{1-r^i}{1-r}\right) & \text{if } p \neq q, \text{ i.e. if } p \neq \frac{1}{2} \\ iP_1 & \text{if } p = \frac{1}{2} \end{cases}$$

Since $P_N = 1$, we have

$$1 = P_N = \begin{cases} P_1\left(\frac{1-r^N}{1-r}\right) & \text{if } p \neq \frac{1}{2} \\ NP_1 & \text{if } p = \frac{1}{2} \end{cases}$$

So $$P_1 = \begin{cases} 1 \cdot \left(\frac{1-r}{1-r^N}\right) & \text{if } p \neq \frac{1}{2} \\ \frac{1}{N} \cdot 1 & \text{if } p = \frac{1}{2} \end{cases}$$

$$\boxed{\therefore P_i = \begin{cases} \frac{1-r^i}{1-r^N} & \text{if } p \neq \frac{1}{2} \\ \frac{i}{N} & \text{if } p = \frac{1}{2} \end{cases}}$$

(d) If a gambler has \$100 and wants to double their money at the roulette table, what is their best bet (probabilistically speaking)?

$$r = \frac{20}{18} \qquad P_{100} = \frac{1 - \left(\frac{20}{18}\right)^{100}}{1 - \left(\frac{20}{18}\right)^{200}} \approx 0.0000266$$

prob. that they double their money if they bet \$1 at a time.

Bet it all on the first bet! → prob. that they double their money is $\frac{18}{38}$

4.6 Mean Time Spent in Transient States

In the gambler's ruin problem, how many games do we expect the gambler to play before they either reach a fortune of N or lose it all? How many times do we expect them to return to the fortune that they started with? Here, we will learn how to answer these types of questions.

Consider a finite-state Markov chain.

- Let T denote the set of transient states, and write $T = \{1, 2, \dots, t\}$.
- Let $P_T = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1t} \\ P_{21} & P_{22} & \cdots & P_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ P_{t1} & P_{t2} & \cdots & P_{tt} \end{bmatrix}$ Submatrix of P corresponding to transient states.
- Note that at least one of the row sums of P_T is less than 1. Why?

  Since these are transient, we must eventually leave and never come back. So there is some non zero prob. that we transition to a recurrent state.

Let $i, j \in T$. That is, suppose that i and j are transient states.

- Let s_{ij} denote the expected number of time periods that the process is in state j, given that it starts in state i.
- Let $\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$

  s_{ij}: start in i; expected # of time periods in j.

Conditioning on the outcome of the first transition, we find

$$s_{ij} = \delta_{ij} + \sum_{\text{all } k} P_{ik} s_{kj} = \delta_{ij} + \sum_{k=1}^{t} P_{ik} s_{kj}$$

since $s_{kj} = 0$ if k is not transient.

- Let $S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1t} \\ s_{21} & s_{22} & \cdots & s_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ s_{t1} & s_{t2} & \cdots & s_{tt} \end{bmatrix}$
- Then in matrix notation, we have

$$S = \underbrace{I}_{\text{identity of size } t} + P_T S$$

$$\Rightarrow S - P_T S = I$$

$$\Rightarrow (I - P_T) S = I$$

$$\Rightarrow S = (I - P_T)^{-1}$$

Proposition 4.6.1. *Consider a finite state Markov chain, and let $T = \{1, 2, \ldots, t\}$ denote the set of transient states, and let P_T denote the submatrix of the transition probability matrix corresponding to the states in T. Let s_{ij} be the mean time spent in transient state j, starting from transient state i. Then the $t \times t$ matrix*

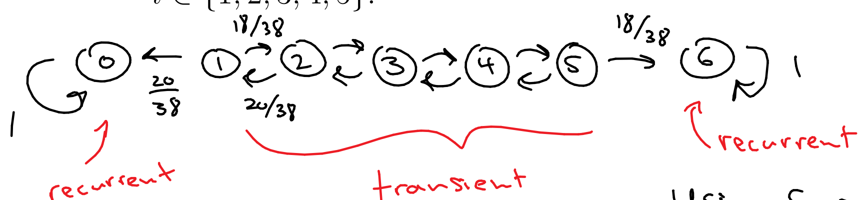
$$S = [s_{ij}]$$

is given by

$$S = (I - P_T)^{-1}$$

Example 4.6.2. Consider the gambler's ruin problem with $N = 6$ and $p = \frac{18}{38}$. Suppose that the gambler starts with three units.

(a) Find the expected number of rounds that the gambler's fortune is i units for all $i \in \{1, 2, 3, 4, 5\}$.



$$P_T = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 0 & 18/38 & 0 & 0 & 0 \\ 20/38 & 0 & 18/38 & 0 & 0 \\ 0 & 20/38 & 0 & \frac{18}{38} & 0 \\ 0 & 0 & \frac{20}{38} & 0 & \frac{18}{38} \\ 0 & 0 & 0 & \frac{20}{38} & 0 \end{bmatrix} \end{array}$$

Using Sage,

expected time in state column #.

e.g. this is s_{33}, the expected # of time periods in state 3, starting from state 3.

$$S = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{bmatrix} & & & & \\ & & & & \\ 1.10 & 2.09 & 2.98 & 1.88 & 0.89 \\ & & & & \\ & & & & \end{bmatrix}$$

starting state (row #)

This is s_{31}, the expected # of time periods spent in state 1, starting from state 3.

(b) Find the expected total number of rounds that the gambler plays before they leave.

The expected total number of rounds is

$$\sum_{j=1}^{5} s_{3j} = 1.10 + 2.09 + 2.98 + 1.88 + 0.89 = 8.93$$

Observation 4.6.3. Let r_i denote the expected total time spent in transient states, starting in transient state i. Then

$$r_i = \sum_{j=1}^{t} s_{ij} = s_{i1} + s_{i2} + \ldots + s_{it}$$

In matrix notation, if $R = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_t \end{bmatrix}$, then $$R = S\vec{1}$$

all 1's vector $(t \times 1)$

Recall that for distinct states i and j, we defined $f_{i,j}$ as the probability that, starting in state i, the process ever visits state j:

$$f_{i,j} = P\{\text{ever visit state } j \mid X_0 = i\}$$

More precisely, we have

$$f_{i,j} = P\{\exists\, n>0 \text{ such that } X_n = j \mid X_0 = i\} = P\left\{ \bigcup_{n=1}^{\infty} \{X_n = j\} \,\middle|\, X_0 = i \right\}$$

Proposition 4.6.4. *For transient states i and j in a finite-state Markov chain, we have*

$$f_{i,j} = s_{ij}/s_{jj}$$

Proof. We find an expression for s_{ij} by conditioning on whether or not we <u>ever</u> enter state j:

$$s_{ij} = E[\text{time in } j \mid \text{start in } i, \text{ ever visit } j] \cdot P\{\text{ever visit } j \mid \text{start in } i\}$$
$$+ E[\text{time in } j \mid \text{start in } i, \text{ never visit } j] \cdot P\{\text{never visit } j \mid \text{start in } i\}$$

$$= s_{jj} \cdot f_{ij} + 0 \cdot (1 - f_{ij})$$

$$\Rightarrow \quad s_{ij} = s_{jj} \cdot f_{ij} \quad \Rightarrow \quad f_{ij} = \frac{s_{ij}}{s_{jj}}$$

$\square$

Example 4.6.5. In the previous example, find the probability that the gambler's fortune ever goes down to 1.

Remember that $X_0 = 3$, i.e., the gambler starts with 3 units. So we want

$$f_{31} = \frac{s_{31}}{s_{11}}$$

$$\simeq \frac{1.1}{1.66} \quad \text{(from Sage)}$$

$$\approx \frac{2}{3}$$

Example 4.6.6. For the gambler's ruin problem with arbitrary $N > 0$ and $p \in (0, 1)$, let M_i denote the mean number of games played until the gambler either goes broke or reaches a fortune of N, given that they start with i units. Find M_i.

Note: We will start this example in class, but you will finish it on your own on the next assignment. See Exercise 59 in the text.

We have

$$M_0 = 0 = M_N$$

For all $i \in \{1, 2, 3, \ldots, N-1\}$, by conditioning on the outcome of the first play, we have.

$$M_i = 1 + p M_{i+1} + q M_{i-1}$$

$q = 1 - p$

4.7 Branching Processes

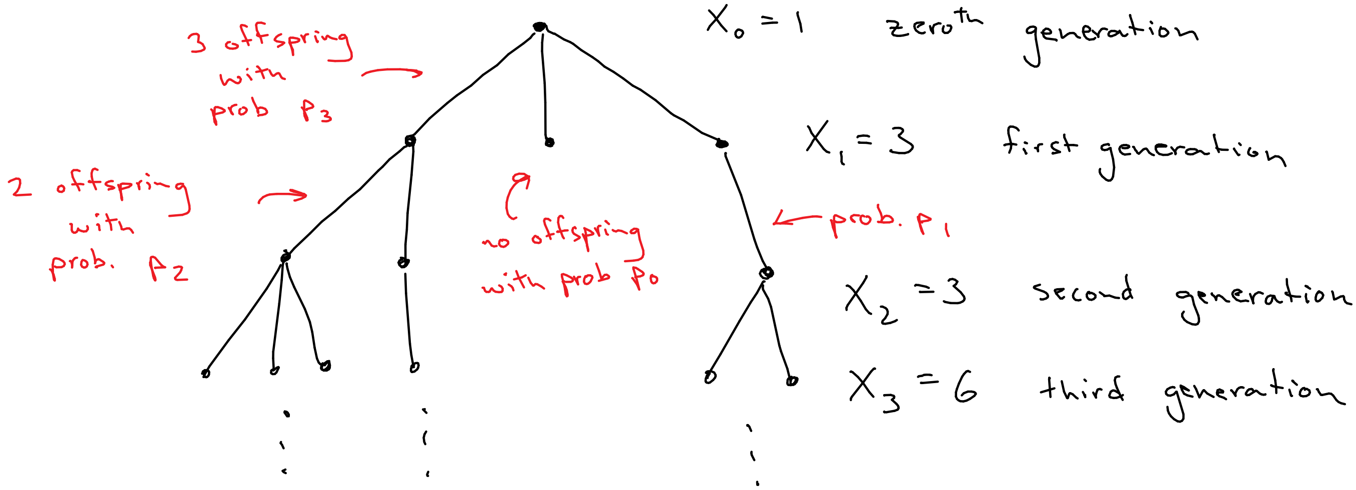
Consider a population consisting of individuals able to produce offspring of the same kind.

- Suppose that each individual will, by the end of its lifetime, produce j new offspring with probability p_j, independently of the reproduction of all others, and the total number of individuals present.

- Let X_0 denote the number of individuals initially present. These individuals make up the *zeroth* generation.

- Let X_1 denote the total number of offspring of members of the zeroth generation. These offspring make up the *first generation*.

- Let X_2 denote the total number of offspring of members of the first generation. These offspring make up the *second generation*. And so on!



Then $\{X_n,\ n = 0, 1, 2, \ldots\}$ is a special type of Markov chain called a *branching process*.

- Note that $P_{00} = 1$, so state 0 is recurrent (as an absorbing state).

- If $p_0 > 0$, then all other states are transient, since for any state $i > 0$, we have

$$P_{i0} = p_0^i > 0 .$$

(So from any state $i > 0$, there is a nonzero prob. that we leave and never come back.)

- Since the process only visits each transient state a finite number of times, it follows that the population will either die out or grow arbitrarily large.

Given the number of individuals initially present and the probability distribution of the number of offspring of each individual, we learn how to answer the following questions.

- What are $E[X_n]$ and $\mathrm{Var}(X_n)$?

- What is the probability that the population eventually dies out?

For the proof of our first proposition, recall that if $S = \sum_{i=1}^{N} Z_i$ is a compound random variable with $E[Z_i] = \mu$ and $\text{Var}(Z_i) = \sigma^2$, then

$$E[S] = \mu E[N] \qquad \text{and } \text{Var}(S) = \mu^2 \text{Var}(N) + \sigma^2 E[N]$$

Proposition 4.7.1. *Let $\{X_n,\ n = 0, 1, 2, \ldots\}$ be a branching process in which $X_0 = 1$, and the number of offspring that each individual produces has mean μ and variance σ^2. Then*

$$E[X_n] = \mu^n$$

and

$$\text{Var}(X_n) = \sigma^2 \left(\mu^{n-1} + \mu^n + \cdots + \mu^{2n-2}\right) = \begin{cases} \sigma^2 \mu^{n-1} \left(\dfrac{1-\mu^n}{1-\mu}\right) & \text{if } \mu \neq 1 \\ n\sigma^2 & \text{if } \mu = 1 \end{cases}$$

Proof. For each $n > 0$, we write

$$X_n = \sum_{i=1}^{X_{n-1}} Z_i,$$

where Z_i is the number of offspring of the ith individual of the $(n-1)$th generation.

- So X_n is a compound random variable!

Since

$$E[Z_i] = \mu \qquad \text{and } \text{Var}(Z_i) = \sigma^2$$

we have

$$E[X_n] = \mu E[X_{n-1}] \qquad \text{and } \text{Var}(X_n) = \mu^2 \text{Var}(X_{n-1}) + \sigma^2 E[X_{n-1}]$$

We now use induction.

- Expectation: We want to show that $E[X_n] = \mu^n$ for all $n \geq 0$.

Base Case: $E[X_0] = 1 = \mu^0$ ✓

Inductive Hypothesis: Suppose for some $n > 0$ that

$$E[X_{n-1}] = \mu^{n-1}.$$

Inductive Step: Then $E[X_n] = \mu E[X_{n-1}] = \mu \cdot \mu^{n-1} = \mu^n$.

$\therefore$ By math induction, $E[X_n] = \mu^n$ for all $n \geq 0$.

$\mathrm{Var}(X_0) = 0$

- Variance: So far, we have shown that for all $n > 0$, we have

$$\mathrm{Var}(X_n) = \mu^2 \, \mathrm{Var}(X_{n-1}) + \sigma^2 E[X_{n-1}].$$

We want to show that for all $n \geq 1$, we have

$$\mathrm{Var}(X_n) = \sigma^2(\mu^{n-1} + \mu^n + \cdots + \mu^{2n-2}).$$

Base Case: $\mathrm{Var}(X_1) = \mu^2 \mathrm{Var}(X_0) + \sigma^2 E[X_0]$

$= 0 + \sigma^2 \cdot 1$

$= \sigma^2$

$2(n-1)-2$

Inductive Hypothesis: Suppose that for some $n > 1$ that

$\mathrm{Var}(X_{n-1}) = \sigma^2(\mu^{n-2} + \mu^{n-1} + \ldots + \mu^{2n-4})$

Inductive Step: We have

$\mathrm{Var}(X_n) = \mu^2 \mathrm{Var}(X_{n-1}) + \sigma^2 E[X_{n-1}]$

$= \mu^2 \sigma^2 (\mu^{n-2} + \mu^{n-1} + \ldots + \mu^{2n-4}) + \sigma^2 \mu^{n-1}$

$= \sigma^2(\mu^n + \mu^{n+1} + \ldots + \mu^{2n-2}) + \sigma^2 \mu^{n-1}$

$= \sigma^2(\mu^{n-1} + \mu^n + \mu^{n+1} + \ldots + \mu^{2n-2})$

$\therefore$ By math induction, $\mathrm{Var}(X_n) = \sigma^2(\mu^{n-1} + \mu^n + \ldots + \mu^{2n-2})$ for all $n \geq 1$.

Let π_0 denote the probability that the population eventually dies out, given that $X_0 = 1$. More formally, we define

$$\pi_0 = P\{\exists n \text{ such that } X_n = 0 \mid X_0 = 1\}.$$

Proposition 4.7.2. *Let $\{X_n,\ n = 0, 1, 2, \ldots\}$ be a branching process with notation as above, and assume that $X_0 = 1$.*

- *If $\mu \leq 1$, then*

$$\pi_0 = 1.$$

- *If $\mu > 1$, then π_0 is the smallest positive solution of the equation*

$$x = \sum_{j=0}^{\infty} x^j \cdot p_j$$

prob. that an individual has j offspring.

Example 4.7.3. Consider a branching process $\{X_n,\ n = 0, 1, 2, \ldots\}$ in which $X_0 = 1$, and the probability mass function for the number of offspring of each individual is given by

$$p_0 = \tfrac{1}{4}, \quad p_1 = \tfrac{1}{3}, \quad p_2 = \tfrac{1}{3}, \quad p_3 = \tfrac{1}{12}.$$

(a) Find $E[X_n]$ and $\mathrm{Var}(X_n)$.

First find the mean and variance of the number of offspring of each individual.

$$\mu = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3 = \frac{1}{3} + \frac{2}{3} + \frac{3}{12}$$

$$= \frac{5}{4}$$

σ^2 can be found in a similar manner.

Therefore,

$$E[X_n] = \mu^n = \left(\frac{5}{4}\right)^n$$

and

$$\mathrm{Var}(X_n) = \sigma^2 \mu^{n-1} \left(\frac{1-\mu^n}{1-\mu}\right)$$

$$= \ldots$$

(b) Find the probability that the population eventually dies out.

We solve

$$x = \sum_{j=0}^{\infty} x^j p_j$$

$$\Rightarrow \quad x = 1 \cdot \frac{1}{4} + x \cdot \frac{1}{3} + x^2 \cdot \frac{1}{3} + x^3 \cdot \frac{1}{12}$$

Using sage, we find $x = 1$ or $x = \dfrac{-5 \pm \sqrt{37}}{2}$

$$\therefore \quad \pi_0 = \frac{-5 + \sqrt{37}}{2} \approx 0.541$$

Think about how your answers would change if we had $X_0 = m$ instead of $X_0 = 1$.