

Chapter 4

Markov Chains

4.1 Definitions and Examples

The importance of Markov chains comes from two facts: (i) there are a large number of physical, biological, economic, and social phenomena that can be described in this way, and (ii) there is a well-developed theory that allows us to do computations. We begin with a famous example, then describe the property that is the defining feature of Markov chains.

Example 4.1. Gambler's ruin. Consider a gambling game in which on any turn you win \$1 with probability $p = 0.4$ or lose \$1 with probability $1 - p = 0.6$. Suppose further that you adopt the rule that you quit playing if your fortune reaches $\$N$. Of course, if your fortune reaches $\$0$ the casino makes you stop.

Let X_n be the amount of money you have after n plays. I claim that your fortune, X_n has the “Markov property.” In words, this means that given the current state, any other information about the past is irrelevant for predicting the next state X_{n+1} . To check this, we note that if you are still playing at time n , i.e., your fortune $X_n = i$ with $0 < i < N$, then for any possible history of your wealth $i_{n-1}, i_{n-2}, \dots, i_1, i_0$

$$P(X_{n+1} = i + 1 | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = 0.4$$

since to increase your wealth by one unit you have to win your next bet and the outcome of the previous bets has no useful information for predicting the next outcome.

Turning now to the formal definition, we say that X_n is a discrete time **Markov chain** with **transition matrix** $p(i, j)$ if for any $j, i, i_{n-1}, \dots, i_0$

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = p(i, j) \quad (4.1)$$

Equation (4.1), also called the “Markov property” says that the conditional probability $X_{n+1} = j$ given the entire history $X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 =$

$i_1, X_0 = i_0$ is the same as the conditional probability $X_{n+1} = j$ given only the previous state $X_n = i$. This is what we mean when we say that “given the current state any other information about the past is irrelevant for predicting X_{n+1} .”

In formulating (4.1) we have restricted our attention to the **temporally homogeneous** case in which the **transition probability**

$$p(i, j) = P(X_{n+1} = j | X_n = i)$$

does not depend on the time n . Intuitively, the transition probability gives the rules of the game. It is the basic information needed to describe a Markov chain. In the case of the gambler’s ruin chain, the transition probability has

$$\begin{aligned} p(i, i+1) &= 0.4, & p(i, i-1) &= 0.6, & \text{if } 0 < i < N \\ p(0, 0) &= 1 & p(N, N) &= 1 \end{aligned}$$

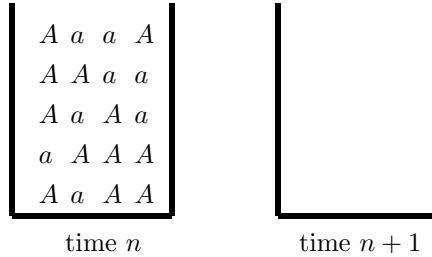
When $N = 5$ the matrix is

	0	1	2	3	4	5
0	1.0	0	0	0	0	0
1	0.6	0	0.4	0	0	0
2	0	0.6	0	0.4	0	0
3	0	0	0.6	0	0.4	0
4	0	0	0	0.6	0	0.4
5	0	0	0	0	0	1.0

Example 4.2. Wright–Fisher model. We consider a fixed population of N genes that can be one of two types: A or a . These types are called alleles. In the simplest version of this model the population at time $n+1$ is obtained by drawing with replacement from the population at time n . In this case if we let X_n be the number of A alleles at time n , then X_n is a Markov chain with transition probability

$$p(i, j) = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j} \quad 0 \leq i, j \leq N$$

since the right-hand side is the binomial distribution for N independent trials with success probability i/N . Note that when $i = 0$, $p(0, 0) = 1$, and when $i = N$, $p(N, N) = 1$.



In the Gambler's Ruin chain and the Wright-Fisher model the states 0 and N are **absorbing states**. Once we enter these states we can never leave. The long run behavior of these models is not very interesting, they will eventually enter one of the absorbing states and stay there forever. To make the Wright-Fisher model more interesting and more realistic, we can introduce the possibility of mutations: an A that is drawn ends up being an a in the next generation with probability u , while an a that is drawn ends up being an A in the next generation with probability v . In this case the probability an A is produced by a given draw is

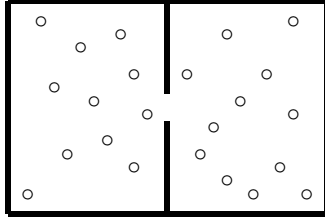
$$\rho_i = \frac{i}{N}(1-u) + \frac{N-i}{N}v$$

i.e., we can get an A by drawing an A and not having a mutation or by drawing an a and having a mutation. Since the draws are independent the transition probability still has the binomial form

$$p(i, j) = \binom{N}{j} (\rho_i)^j (1 - \rho_i)^{N-j} \quad (4.2)$$

Moving from biology to physics:

Example 4.3. Ehrenfest chain. We imagine two cubical volumes of air connected by a small hole. In the mathematical version, we have two “urns,” i.e., two of the exalted trash cans of probability theory, in which there are a total of N balls. We pick one of the N balls at random and move it to the other urn.



Let X_n be the number of balls in the “left” urn after the n th draw. It should be clear that X_n has the Markov property; i.e., if we want to guess the state at time $n + 1$, then the current number of balls in the left urn X_n , is the only relevant information from the observed sequence of states $X_n, X_{n-1}, \dots, X_1, X_0$. To check this we note that

$$P(X_{n+1} = i + 1 | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \frac{N - i}{N}$$

since to increase the number we have to pick one of the $N - i$ balls in the other urn. The number can also decrease by 1 with probability i/N . In symbols, we

have computed that the transition probability is given by

$$p(i, i+1) = (N-i)/N, \quad p(i, i-1) = \frac{i}{N} \quad \text{for } 0 \leq i \leq N$$

with $p(i, j) = 0$ otherwise. When $N = 5$, for example, the matrix is

	0	1	2	3	4	5
0	0	5/5	0	0	0	0
1	1/5	0	4/5	0	0	0
2	0	2/5	0	3/5	0	0
3	0	0	3/5	0	2/5	0
4	0	0	0	4/5	0	1/5
5	0	0	0	0	5/5	0

Here we have written 1 as 5/5 to emphasize the pattern in the diagonals of the matrix.

Moving from science to business:

Example 4.4. Inventory chain. An electronics store sells a video game system. If at the end of the day, the number of units they have on hand is 1 or 0, they order enough new units so their total on hand is 5. This chain is an example of an s, S inventory control policy with $s = 1$ and $S = 5$. That is, when the stock on hand falls to s or below we order enough to bring it back up to S .

For simplicity we assume that the new merchandise arrives before the store opens the next day. Let X_n be the number of units on hand at the end of the n th day. If we assume that the number of customers who want to buy a video game system each day is 0, 1, 2, or 3 with probabilities .3, .4, .2, and .1, then we have the following transition matrix:

	0	1	2	3	4	5
0	0	0	.1	.2	.4	.3
1	0	0	.1	.2	.4	.3
2	.3	.4	.3	0	0	0
3	.1	.2	.4	.3	0	0
4	0	.1	.2	.4	.3	0
5	0	0	.1	.2	.4	.3

To explain the entries we note that if $X_n \geq 2$ then no ordering is done so what we have at the end of the day is the supply minus the demand. If $X_n = 2$ and the demand is 3 or more, or if $X_n = 3$ and the demand is 4, we end up with 0 units at the end of the day and at least one unhappy customer. If $X_n = 0$ or 1 then we will order enough so that at the beginning of the day we have 5, so the result at the end of the day is the same as if $X_n = 5$.

Markov chains are described by giving their transition probabilities. To create a chain, we can write down any $n \times n$ matrix, provided that the entries satisfy:

- (i) $p(i, j) \geq 0$, since they are probabilities.
- (ii) $\sum_j p(i, j) = 1$, since when $X_n = i$, X_{n+1} will be in some state j .

The equation in (ii) is read “sum $p(i, j)$ over all possible values of j .” In words the last two conditions say: the entries of the matrix are nonnegative and each **row** of the matrix sums to 1.

Any matrix with properties (i) and (ii) gives rise to a Markov chain, X_n . To construct the chain we can think of playing a board game. When we are in state i , we roll a die (or generate a random number on a computer) to pick the next state, going to j with probability $p(i, j)$. To illustrate this we will now introduce some simple examples.

Example 4.5. Weather chain. Let X_n be the weather on day n in Ithaca, NY, which we assume is either: 1 = *rainy*, or 2 = *sunny*. Even though the weather is not exactly a Markov chain, we can propose a Markov chain model for the weather by writing down a transition probability

	1	2
1	.6	.4
2	.2	.8

The table says, for example, the probability a rainy day (state 1) is followed by a sunny day (state 2) is $p(1, 2) = 0.6$.

Example 4.6. Social mobility. Let X_n be a family’s social class in the n th generation, which we assume is either 1 = *lower*, 2 = *middle*, or 3 = *upper*. In our simple version of sociology, changes of status are a Markov chain with the following transition probability

	1	2	3
1	.7	.2	.1
2	.3	.5	.2
3	.2	.4	.4

Example 4.7. Brand preference. Suppose there are three types of laundry detergent, 1, 2, and 3, and let X_n be the brand chosen on the n th purchase. Customers who try these brands are satisfied and choose the same thing again with probabilities 0.8, 0.6, and 0.4 respectively. When they change they pick one of the other two brands at random. The transition probability is

	1	2	3
1	.8	.1	.1
2	.2	.6	.2
3	.3	.3	.4

Example 4.8. Two-stage Markov chains. In a Markov chain the distribution of X_{n+1} only depends on X_n . This can easily be generalized to case in which the distribution of X_{n+1} only depends on (X_n, X_{n-1}) . For a concrete example consider a basketball player who makes a shot with the following probabilities:

- 1/2 if he has missed the last two times
- 2/3 if he has hit one of his last two shots
- 3/4 if he has hit both of his last two shots

To formulate a Markov chain to model his shooting, we let the states of the process be the outcomes of his last two shots: $\{HH, HM, MH, MM\}$ where M is short for miss and H for hit. The transition probability is

	<i>HH</i>	<i>HM</i>	<i>MH</i>	<i>MM</i>
<i>HH</i>	3/4	1/4	0	0
<i>HM</i>	0	0	2/3	1/3
<i>MH</i>	2/3	1/3	0	0
<i>MM</i>	0	0	1/2	1/2

To explain suppose the state is HM , i.e., $X_{n-1} = H$ and $X_n = M$. In this case the next outcome will be H with probability 2/3. When this occurs the next state will be $(X_n, X_{n+1}) = (M, H)$ with probability 2/3. If he misses an event of probability 1/3, $(X_n, X_{n+1}) = (M, M)$.

The Hot Hand is a phenomenon known to everyone who plays or watches basketball. After making a couple of shots, players are thought to “get into a groove” so that subsequent successes are more likely. Purvis Short of the Golden State Warriors describes this more poetically as

“You’re in a world all your own. It’s hard to describe. But the basket seems to be so wide. No matter what you do, you know the ball is going to go in.”

Unfortunately for basketball players, data collected by Tversky and Gliovich (*Chance* vol. 2 (1989), No. 1, pages 16–21) shows that this is a misconception. The next table gives data for the conditional probability of hitting a shot after missing the last three, missing the last two, . . . hitting the last three, for nine players of the Philadelphia 76ers: Darryl Dawkins (403), Maurice Cheeks (339), Steve Mix (351), Bobby Jones (433), Clint Richardson (248), Julius Erving (884), Andrew Toney (451), Caldwell Jones (272), and Lionel Hollins (419). The numbers in parentheses are the number of shots for each player.

$P(H 3M)$	$P(H 2M)$	$P(H 1M)$	$P(H 1H)$	$P(H 2H)$	$P(H 3H)$
.88	.73	.71	.57	.58	.51
.77	.60	.60	.55	.54	.59
.70	.56	.52	.51	.48	.36
.61	.58	.58	.53	.47	.53
.52	.51	.51	.53	.52	.48
.50	.47	.56	.49	.50	.48
.50	.48	.47	.45	.43	.27
.52	.53	.51	.43	.40	.34
.50	.49	.46	.46	.46	.32

In fact, the data supports the opposite assertion: after missing a player is more conservative about the shots that they take and will hit more frequently.

4.2 Multistep Transition Probabilities

The previous section introduced several examples to think about. The basic question concerning Markov chains is what happens in the long run? In the case of the weather chain, does the probability that day n is sunny converge to a limit? In the case of the social mobility and brand preference chain, do the fractions of the population in the three income classes (or that buy each of the three types of detergent) stabilize as time goes on? The first step in answering these questions is to figure out what happens in the Markov chain after two or more steps.

The transition probability $p(i, j) = P(X_{n+1} = j | X_n = i)$ gives the probability of going from i to j in one step. Our goal in this section is to compute the probability of going from i to j in $m > 1$ steps:

$$p^m(i, j) = P(X_{n+m} = j | X_n = i)$$

For a concrete example, we start with the transition probability of the social mobility chain:

	1	2	3
1	.7	.2	.1
2	.3	.5	.2
3	.2	.4	.4

To warm-up we consider:

Example 4.9. Suppose the family starts in the middle class (state 2) in generation 0. What is the probability that the generation 1 rises to the upper class (state 3) and generation 2 falls to the lower class (state 1)?

Intuitively, the Markov property implies that starting from state 2 the probability of jumping to 1 and then to 3 is given by $p(2, 3)p(3, 1)$. To get this conclusion from the definitions, we note that using the definition of conditional probability,

$$P(X_2 = 1, X_1 = 3 | X_0 = 2) = \frac{P(X_2 = 1, X_1 = 3, X_0 = 2)}{P(X_0 = 2)}$$

Multiplying and dividing by $P(X_1 = 3, X_0 = 2)$:

$$= \frac{P(X_2 = 1, X_1 = 3, X_0 = 2)}{P(X_1 = 3, X_0 = 2)} \cdot \frac{P(X_1 = 3, X_0 = 2)}{P(X_0 = 2)}$$

Using the definition of conditional probability:

$$= P(X_2 = 1 | X_1 = 3, X_0 = 2) \cdot P(X_1 = 3 | X_0 = 2)$$

By the Markov property (4.1) the last expression is

$$P(X_2 = 1 | X_1 = 3) \cdot P(X_1 = 3 | X_0 = 2) = p(2, 3)p(3, 1) \quad \square$$

Moving on to the real question:

Example 4.10. Suppose the family starts in the middle class (state 2) in generation 0. What is the probability that generation 2 will be in the lower class (state 1)?

To do this we simply have to consider the three possible states for generation 1 and use the previous computation.

$$\begin{aligned} P(X_2 = 1|X_0 = 2) &= \sum_{k=1}^3 p(2, k)p(k, 1) \\ &= (.3)(.7) + (.5)(.3) + (.2)(.2) \\ &= .21 + .15 + .04 = .40 \end{aligned}$$

There is nothing special here about the states 2 and 1 here. By the same reasoning,

$$P(X_2 = j|X_0 = i) = \sum_{k=1}^3 p(i, k)p(k, j)$$

The right-hand side of the last equation gives the (i, j) th entry of the matrix p is multiplied by itself.

To explain this, we note that to compute $p^2(2, 1)$ we multiplied the entries of the second row by those in the first column:

$$\begin{pmatrix} . & . & . \\ .3 & .5 & .2 \\ . & . & . \end{pmatrix} \begin{pmatrix} .7 & . & . \\ .3 & . & . \\ .2 & . & . \end{pmatrix} = \begin{pmatrix} . & . & . \\ .40 & . & . \\ . & . & . \end{pmatrix}$$

If we wanted $p^2(1, 3)$ we would multiply the first row by the third column:

$$\begin{pmatrix} .7 & .2 & .1 \\ . & . & . \\ . & . & . \end{pmatrix} \begin{pmatrix} . & . & .1 \\ . & . & .2 \\ . & . & .4 \end{pmatrix} = \begin{pmatrix} . & . & .15 \\ . & . & . \\ . & . & . \end{pmatrix}$$

When all of the computations are done we have

$$\begin{pmatrix} .7 & .2 & .1 \\ .3 & .5 & .2 \\ .2 & .4 & .4 \end{pmatrix} \begin{pmatrix} .7 & .2 & .1 \\ .3 & .5 & .2 \\ .2 & .4 & .4 \end{pmatrix} = \begin{pmatrix} .57 & .28 & .15 \\ .40 & .39 & .21 \\ .34 & .40 & .26 \end{pmatrix}$$

The two step transition probability $p^2 = p \cdot p$. Based on this you can probably leap to the next conclusion:

Theorem 4.1. *The m -step transition probability*

$$p^m(i, j) = P(X_{n+m} = j|X_n = i) \quad (4.3)$$

is the m th power of the transition matrix p , i.e., $p \cdot p \cdots p$, where there are m terms in the product.

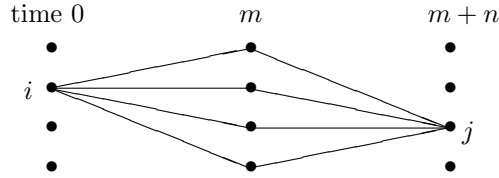
The key ingredient in proving this is the:

Chapman–Kolmogorov equation

$$p^{m+n}(i, j) = \sum_k p^m(i, k) p^n(k, j) \quad (4.4)$$

Once this is proved, (4.3) follows, since taking $n = 1$ in (4.4), we see that $p^{m+1} = p^m \cdot p$.

Why is (4.4) true? To go from i to j in $m + n$ steps, we have to go from i to some state k in m steps and then from k to j in n steps. The Markov property implies that the two parts of our journey are independent. \square



Proof of (4.4). The *independence* in the second sentence of the previous explanation is the mysterious part. To show this, we combine Examples 4.9 and 4.10. Breaking things down according to the state at time m ,

$$P(X_{m+n} = j | X_0 = i) = \sum_k P(X_{m+n} = j, X_m = k | X_0 = i)$$

Repeating the computation in Example 4.9, the definition of conditional probability implies:

$$P(X_{m+n} = j, X_m = k | X_0 = i) = \frac{P(X_{m+n} = j, X_m = k, X_0 = i)}{P(X_0 = i)}$$

Multiplying and dividing by $P(X_m = k, X_0 = i)$ gives:

$$= \frac{P(X_{m+n} = j, X_m = k, X_0 = i)}{P(X_m = k, X_0 = i)} \cdot \frac{P(X_m = k, X_0 = i)}{P(X_0 = i)}$$

Using the definition of conditional probability we have:

$$= P(X_{m+n} = j | X_m = k, X_0 = i) \cdot P(X_m = k | X_0 = i)$$

By the Markov property (4.1) the last expression is

$$= P(X_{m+n} = j | X_m = k) \cdot P(X_m = k | X_0 = i) = p^m(i, k) p^n(k, j)$$

and we have proved (4.4). \square

Having established (4.4), we now return to computations. We begin with the weather chain

$$\begin{pmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix} \begin{pmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix} = \begin{pmatrix} 0.44 & 0.56 \\ 0.28 & 0.72 \end{pmatrix}$$

Multiplying again $p^2 \cdot p = p^3$

$$\begin{pmatrix} 0.44 & 0.56 \\ 0.28 & 0.72 \end{pmatrix} \begin{pmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix} = \begin{pmatrix} 0.376 & 0.624 \\ 0.312 & 0.688 \end{pmatrix}$$

and then $p^3 \cdot p = p^4$

$$\begin{pmatrix} 0.376 & 0.624 \\ 0.312 & 0.688 \end{pmatrix} \begin{pmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix} = \begin{pmatrix} 0.3504 & 0.6496 \\ 0.3248 & 0.6752 \end{pmatrix}$$

To increase the time faster we can use (4.4) to conclude that $p^4 \cdot p^4 = p^8$:

$$\begin{pmatrix} 0.3504 & 0.6496 \\ 0.3248 & 0.6752 \end{pmatrix} \begin{pmatrix} 0.3504 & 0.6496 \\ 0.3248 & 0.6752 \end{pmatrix} = \begin{pmatrix} 0.33377 & 0.66623 \\ 0.33311 & 0.66689 \end{pmatrix}$$

Multiplying again $p^8 \cdot p^8 = p^{16}$

$$\begin{pmatrix} 0.33333361 & 0.66666689 \\ 0.33333319 & 0.66666681 \end{pmatrix}$$

Based on the last calculation, one might guess that as n gets large the matrix becomes closer and closer to

$$\begin{pmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{pmatrix}$$

This is true and will be explained in the next section.

4.3 Stationary distributions

Our first step is to consider

What happens when the initial state is random? Breaking things down according to the value of the initial state and using the definition of conditional probability

$$\begin{aligned} P(X_n = j) &= \sum_i P(X_0 = i, X_n = j) \\ &= \sum_i P(X_0 = i)P(X_n = j|X_0 = i) \end{aligned}$$

If we introduce $q(i) = P(X_0 = i)$, then the last equation can be written as

$$P(X_n = j) = \sum_i q(i)p^n(i, j) \quad (4.5)$$

In words, we multiply the transition matrix on the left by the vector q of initial probabilities. If there are k states, then $p^n(x, y)$ is a $k \times k$ matrix. So to make the matrix multiplication work out right, we should take q as a $1 \times k$ matrix or a “row vector.”

For a concrete example consider the weather chain (Example 4.5) and suppose that the initial distribution is $q(1) = 0.3$ and $q(2) = 0.7$. In this case

$$(0.3 \quad 0.7) \begin{pmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix} = (0.32 \quad 0.68)$$

since

$$\begin{aligned} 0.3(0.6) + 0.7(0.2) &= 0.32 \\ 0.3(0.4) + 0.7(0.8) &= 0.68 \end{aligned}$$

For a second example consider the social mobility chain (Example 4.6) and suppose that the initial distribution: $q(1) = .5$, $q(2) = .2$, and $q(3) = .3$. Multiplying the vector q by the transition probability gives the vector of probabilities at time 1.

$$(.5 \quad .2 \quad .3) \begin{pmatrix} .7 & .2 & .1 \\ .3 & .5 & .2 \\ .2 & .4 & .4 \end{pmatrix} = (.47 \quad .32 \quad .21)$$

To check the arithmetic note that the three entries on the right-hand side are

$$\begin{aligned} .5(.7) + .2(.3) + .3(.2) &= .35 + .06 + .06 = .47 \\ .5(.2) + .2(.5) + .3(.4) &= .10 + .10 + .12 = .32 \\ .5(.1) + .2(.2) + .3(.4) &= .05 + .04 + .12 = .21 \end{aligned}$$

If the distribution at time 0 is the same as the distribution at time 1, then by the Markov property it will be distribution at all times $n \geq 1$. Because of this

q is called a **stationary distribution**. Stationary distributions have a special importance in the theory of Markov chains, so we will use a special letter π to denote solutions of the equation

$$\pi \cdot p = \pi.$$

To have a mental picture of what happens to the distribution of probability when one step of the Markov chain is taken, it is useful to think that we have $q(i)$ pounds of sand at state i , with the total amount of sand $\sum_i q(i)$ being one pound. When a step is taken in the Markov chain, a fraction $p(i, j)$ of the sand at i is moved to j . The distribution of sand when this has been done is

$$q \cdot p = \sum_i q(i)p(i, j)$$

If the distribution of sand is not changed by this procedure q is a stationary distribution.

General two state transition probability.

$$\begin{array}{cc} & \mathbf{1} & \mathbf{2} \\ \mathbf{1} & 1-a & a \\ \mathbf{2} & b & 1-b \end{array}$$

We have written the chain in this way so the stationary distribution has a simple formula

$$\pi(1) = \frac{b}{a+b} \quad \pi(2) = \frac{a}{a+b} \quad (4.6)$$

As a first check on this formula we note that in the weather chain $a = 0.4$ and $b = 0.2$ which gives $(1/3, 2/3)$ as we found before. We can prove this works in general by drawing a picture:

$$\frac{b}{a+b} \mathbf{1} \bullet \begin{array}{c} \xrightarrow{a} \\ \xleftarrow{b} \end{array} \bullet \frac{a}{a+b} \mathbf{2}$$

In words, the amount of sand that flows from 1 to 2 is the same as the amount that flows from 2 to 1 so the amount of sand at each site stays constant. To check algebraically that this is true:

$$\begin{aligned} \frac{b}{a+b}(1-a) + \frac{a}{a+b}b &= \frac{b-ba+ab}{a+b} = \frac{b}{a+b} \\ \frac{b}{a+b}a + \frac{a}{a+b}(1-b) &= \frac{ba+a-ab}{a+b} = \frac{a}{a+b} \end{aligned} \quad (4.7)$$

Formula (4.6) gives the stationary distribution for any two state chain, so we progress now to the three state case and consider the brand preference chain (Example 4.7). The equation $\pi p = \pi$ says

$$(\pi_1 \quad \pi_2 \quad \pi_3) \begin{pmatrix} .8 & .1 & .1 \\ .2 & .6 & .2 \\ .3 & .3 & .4 \end{pmatrix} = (\pi_1 \quad \pi_2 \quad \pi_3)$$

which translates into three equations

$$\begin{aligned} .8\pi_1 + .2\pi_2 + .3\pi_3 &= \pi_1 \\ .1\pi_1 + .6\pi_2 + .3\pi_3 &= \pi_2 \\ .1\pi_1 + .2\pi_2 + .4\pi_3 &= \pi_3 \end{aligned}$$

Note that the columns of the matrix give the numbers in the rows of the equations. The third equation is redundant since if we add up the three equations we get

$$\pi_1 + \pi_2 + \pi_3 = \pi_1 + \pi_2 + \pi_3$$

If we replace the third equation by $\pi_1 + \pi_2 + \pi_3 = 1$ and subtract π_1 from each side of the first equation and π_2 from each side of the second equation we get

$$\begin{aligned} -.2\pi_1 + .2\pi_2 + .3\pi_3 &= 0 \\ .1\pi_1 - .4\pi_2 + .3\pi_3 &= 0 \\ \pi_1 + \pi_2 + \pi_3 &= 1 \end{aligned} \tag{4.8}$$

At this point we can solve the equations by hand or using a calculator.

By hand. We note that the third equation implies $\pi_3 = 1 - \pi_1 - \pi_2$ and substituting this in the first two gives

$$\begin{aligned} -.5\pi_1 - .1\pi_2 + .3 &= 0 \\ -.2\pi_1 - .7\pi_2 + .3 &= 0 \end{aligned}$$

which we rearrange to give

$$\begin{aligned} 0.3 &= .5\pi_1 + .1\pi_2 \\ 0.3 &= .2\pi_1 + .7\pi_2 \end{aligned}$$

Multiplying the first equation by .7 and adding $-.1$ times the second gives

$$1.8 = (.35 - .02)\pi_1 \quad \text{or} \quad \pi_1 = 18/33 = 6/11$$

Multiplying the first equation by .2 and adding $-.5$ times the second gives

$$-0.09 = (0.02 - .35)\pi_2 \quad \text{or} \quad \pi_2 = 9/33 = 3/11$$

Since the three probabilities add up to 1, $\pi_3 = 2/11$.

Using the TI83 calculator is easier. To begin we write (4.8) in matrix form as

$$(\pi_1 \quad \pi_2 \quad \pi_3) \begin{pmatrix} -.2 & .1 & 1 \\ .2 & -.4 & 1 \\ .3 & .3 & 1 \end{pmatrix} = (0 \quad 0 \quad 1)$$

If we let A be the 3×3 matrix in the middle this can be written as $\pi A = (0, 0, 1)$. Multiplying on each side by A^{-1} we see that

$$\pi = (0, 0, 1)A^{-1}$$

which is the third row of A^{-1} . Entering A into our calculator computing the inverse and reading the third row we find that the stationary distribution is

$$(.545454, .272727, .181818)$$

Converting the answer to fractions using the first entry in the math menu gives

$$(6/11, 3/11, 2/11)$$

Example 4.11. Social Mobility (continuation of 4.6).

	1	2	3
1	.7	.2	.1
2	.3	.5	.2
3	.2	.4	.4

Using the first two equations and the fact that the sum of the π 's is 1

$$\begin{aligned} .7\pi_1 + .3\pi_2 + .2\pi_3 &= \pi_1 \\ .2\pi_1 + .5\pi_2 + .4\pi_3 &= \pi_2 \\ \pi_1 + \pi_2 + \pi_3 &= 1 \end{aligned}$$

This translates into $\pi A = (0, 0, 1)$ with

$$A = \begin{pmatrix} -.3 & .2 & 1 \\ .3 & -.5 & 1 \\ .2 & .4 & 1 \end{pmatrix}$$

Note that here and in the previous example the first two columns of A consist of the first two columns of the transition probability with 1 subtracted from the diagonal entries, and the final column is all 1's. Computing the inverse and reading the last row gives

$$(.468085, .340425, .191489)$$

Converting the answer to fractions using the first entry in the math menu gives

$$(22/47, 16/47, 9/47)$$

Example 4.12. Basketball (continuation of 4.8).

To find the stationary matrix in this case we can follow the same procedure. A consists of the first three columns of the transition matrix with 1 subtracted from the diagonal, and a final column of all 1's.

$$\begin{array}{cccc} -1/4 & 1/4 & 0 & 1 \\ 0 & -1 & 2/3 & 1 \\ 2/3 & 1/3 & -1 & 1 \\ 0 & 0 & 1/2 & 1 \end{array}$$

The answer is given by the fourth row of A^{-1} :

$$(.5, .1875, .1875, .125)$$

Thus the long run fraction of time the player hits a shot is $\pi(HH) + \pi(MH) = 0.6875$.

Example 4.13. Inventory chain (continuation of 4.4).

As in the two previous examples, A consists of the first five columns of the transition matrix with 1 subtracted from the diagonal, and a final column of all 1's.

$$\begin{array}{cccccc} -1 & 0 & .1 & .2 & .4 & 1 \\ 0 & -1 & .1 & .2 & .4 & 1 \\ .3 & .4 & -.7 & 0 & 0 & 1 \\ .1 & .2 & .4 & -.7 & 0 & 1 \\ 0 & .1 & .2 & .4 & -.7 & 1 \\ 0 & 0 & .1 & .2 & .4 & 1 \end{array}$$

The answer is given by the sixth row of A^{-1} :

$$(.090862, .155646, .231006, .215605, .201232, .105646)$$

Converting the answer to fractions using the first entry in the math menu gives

$$(177/1948, 379/2435, 225/974, 105/487, 98/487, 1029/9740)$$

but the decimal version is probably more informative.

Example 4.14. Ehrenfest chain (continuation of 4.3).

Consider first the case $N = 5$. As in the three previous examples, A consists of the first five columns of the transition matrix with 1 subtracted from the diagonal, and a final column of all 1's.

$$\begin{array}{cccccc} -1 & 1 & 0 & 0 & 0 & 1 \\ .2 & -1 & .8 & 0 & 0 & 1 \\ 0 & .4 & -1 & .6 & 0 & 1 \\ 0 & 0 & .6 & -1 & .4 & 1 \\ 0 & 0 & 0 & .8 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}$$

The answer is given by the sixth row of A^{-1} :

$$(.03125, .15625, .3125, .3125, .15625, .03125)$$

Even without a calculator we can recognize these as

$$(1/32, 5/32, 10/32, 10/32, 5/32, 1/32)$$

the probabilities of 0 to 5 heads when we flip five coins.

Based on this we can guess that in general

$$\pi(k) = \binom{N}{k} / 2^N$$

Proof by computation. For $0 < k < N$ we can end up in state k only by coming up from $k - 1$ or down from $k + 1$ so

$$\begin{aligned} & \pi(k-1)p(k-1, k) + \pi(k+1)p(k+1, k) \\ &= \frac{\binom{N}{k-1}}{2^N} \cdot \frac{N-k+1}{N} + \frac{\binom{N}{k+1}}{2^N} \cdot \frac{k+1}{N} \\ &= \frac{1}{2^N} \left(\frac{(N-1)!}{(k-1)!(N-k)!} + \frac{(N-1)!}{(k)!(N-k+1)!} \right) \\ &= \frac{1}{2^N} \binom{N}{k} \left(\frac{k}{N} + \frac{N-k}{N} \right) = \pi(k) \end{aligned}$$

The only way to end up at 0 is by coming down from 1 so

$$\pi(1)p(1, 0) = \frac{N}{2^N} \cdot \frac{1}{N} = \pi(0)$$

Similarly, the only way to end up at N is by coming up from $N - 1$ so

$$\pi(N-1)p(N-1, N) = \frac{N}{2^N} \cdot \frac{1}{N} = \pi(0) \quad \square$$

Proof by thinking. To determine the initial state, (a) flip N coins, with heads = in the left urn, and tails = in the right. A transition of the chain corresponds to (b) picking a coin at random and turning it over. It is clear that the end result of (a) and (b) has all 2^N outcomes equally likely, so the state at time 1 is the same as the state at time 0. \square

4.4 Limit Behavior

In this section we will give conditions that guarantee that as n gets large $p^n(i, j)$ approaches its stationary distribution. We begin with the

Convergence in the two state case. Let p_0 be the initial probability of being in state 1, and let p_n be the probability of being in state 1 after n steps.

Theorem 4.2. *For a two state Markov chain with transition probability*

$$\begin{array}{cc} 1-a & a \\ b & 1-b \end{array}$$

where $0 < a + b < 2$, we have

$$|p_n - b/(a + b)| = |p_0 - b/(a + b)| \cdot |1 - a - b|^n \quad (4.9)$$

In words, the transition probability converges to equilibrium exponentially fast. In the case of the weather chain $|1 - a - b| = 0.4$, so the difference between p_n and the limit $b/(a + b)$ goes to 0 faster than $(0.4)^n$.

Proof. Using the Markov property we have for any $n \geq 1$ that

$$p_n = p_{n-1}(1 - a) + (1 - p_{n-1})b$$

In words, the chain is in state 1 at time n if it was in state 1 at time $n - 1$ (with probability p_{n-1}) and stays there (with probability $1 - a$), or if it was in state 2 (with probability $1 - p_{n-1}$) and jumps from 2 to 1 (with probability b). Since the probability of being in state 1 is constant when we start in the stationary distribution, see the first equation in (4.7):

$$\frac{b}{a + b} = \frac{b}{a + b}(1 - a) + \left(1 - \frac{b}{a + b}\right)b$$

Subtracting this equation from the one for p_n we have

$$\begin{aligned} p_n - \frac{b}{a + b} &= \left(p_{n-1} - \frac{b}{a + b}\right)(1 - a) + \left(\frac{b}{a + b} - p_{n-1}\right)b \\ &= \left(p_{n-1} - \frac{b}{a + b}\right)(1 - a - b) \end{aligned}$$

If $0 < a + b < 2$ then $|1 - a - b| < 1$ and we have

$$\left|p_n - \frac{b}{a + b}\right| = \left|p_{n-1} - \frac{b}{a + b}\right| \cdot |1 - a - b|$$

In words, the difference $|p_n - b/(a + b)|$ will shrink by a factor $|1 - a - b|$ at each step. Iterating the last equation gives the desired result. \square

There are two cases $a = b = 0$ and $a = b = 1$ in which $p^n(i, j)$ does not converge to $\pi(i)$. In the first case the matrix is

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

so the state never changes. This is called the **identity matrix** and denoted by I , since for any 2×2 matrix m , $I \cdot m = m$ and $m \cdot I = m$. In the second case the matrix is

$$p = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

so the chain always jumps. In this case $p^2 = I$, $p^3 = p$, $p^4 = I$, etc. To see that something similar can happen in a “real example.”

Example 4.15. Ehrenfest chain. Consider the chain defined in Example 4.3 and for simplicity, suppose there are three balls. In this case the transition probability is

	0	1	2	3
0	0	3/3	0	0
1	1/3	0	2/3	0
2	0	2/3	0	1/3
3	0	0	3/3	0

In the second power of p the zero pattern is shifted:

	0	1	2	3
0	1/3	0	2/3	0
1	0	7/9	0	2/9
2	2/9	0	7/9	0
3	0	2/3	0	1/3

To see that the zeros will persist, note that if initially we have an odd number of balls in the left urn, then no matter whether we add or subtract one ball the result will be an even number. Thus X_n alternates between being odd and even. To see why this prevents convergence note that $p^{2n}(i, i) > 0$ while $p^{2n+1}(i, i) = 0$.

A second thing that can prevent convergence is shown by

Example 4.16. Reducible chain.

	0	1	2	3
0	2/3	1/3	0	0
1	1/5	4/5	0	0
2	0	0	1/2	1/2
3	0	0	1/6	5/6

In this case if we start at 0 or 1 it is impossible to get to states 2 or 3 and vice versa, so the 2×2 blocks of 0's will persist forever in the matrix.

A remarkable fact about Markov chains (on finite state spaces) is that if we avoid these two problems then there is a unique stationary distribution π and $p^n(i, j) \rightarrow \pi(j)$. The two conditions are

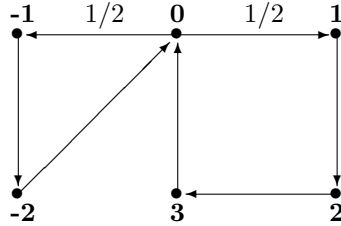
- p is said to be **irreducible** if for each i and j it is possible to get from i to j , i.e., $p^m(i, j) > 0$ for some $m \geq 1$.
- a state i is said to be **aperiodic** if the greatest common divisor of $J_i = \{n \geq 1 \text{ that have } p^n(i, i) > 0\}$ is 1.

In general the greatest common divisor of J_i is called the **period** of state i . In the Ehrenfest chain it is only possible to go from i to i in an even number of steps so all states have period 2. The next example explains why the definition is formulated in terms of the greatest common divisor.

Example 4.17. Triangle and Square. The state space is $\{-2, -1, 0, 1, 2, 3\}$ and the transition probability is

	-2	-1	0	1	2	3
-2	0	0	1	0	0	0
-1	1	0	0	0	0	0
0	0	1/2	0	1/2	0	0
1	0	0	0	0	1	0
2	0	0	0	0	0	1
3	0	0	1	0	0	0

In words, from 0 we are equally likely to go to 1 or -1 . From -1 we go with probability one to -2 and then back to 0, from 1 we go to 2 then to 3 and back to 0. The name refers to the fact that $-1 \rightarrow -2 \rightarrow 0$ is a triangle and $1 \rightarrow 2 \rightarrow 3 \rightarrow 0$ is a square.



In this case it is easy to check that

$$J_0 = \{3, 4, 6, 7, 9, 10, 11, 12, \dots\}$$

so the greatest common divisor of J_0 is 1. In this case and in general for aperiodic states J_0 contains all integers beyond some point.

With the key definitions made we can now state the

Convergence Theorem. *If p is irreducible and has an aperiodic state then there is a unique stationary distribution π and for any i and j*

$$p^n(i, j) \rightarrow \pi(j) \quad \text{as } n \rightarrow \infty \quad (4.10)$$

An easy, but important, special case is

Corollary. *If for some n $p^n(i, j) > 0$ for all i and j then there is a unique stationary distribution π and*

$$p^n(i, j) \rightarrow \pi(j) \quad \text{as } n \rightarrow \infty \quad (4.11)$$

Proof. In this case p is irreducible since it is possible to get from any state to any other in n steps. All states are aperiodic since we also have $p^{n+1}(i, j) > 0$, so $n, n+1 \in J_i$ and hence the greatest common divisor of all the numbers in J_i is 1. \square

The Corollary with $n = 1$ shows that the convergence theorem applies to the Wright-Fisher model with mutation, weather chain, social mobility, and brand preference chains that are Examples 4.2, 4.5, 4.6, and 4.7. The Convergence Theorem does not apply to the Gambler's Ruin chain (Example 4.1) or the Wright-Fisher model with no mutations since they have absorbing states and hence are not irreducible. We have already noted that the Ehrenfest chain (Example 4.3) does not converge since all states have period 2. This leaves the inventory chain (Example 4.4):

0	1	2	3	4	5	
0	0	0	.1	.2	.4	.3
1	0	0	.1	.2	.4	.3
2	.3	.4	.3	0	0	0
3	.1	.2	.4	.3	0	0
4	0	.1	.2	.4	.3	0
5	0	0	.1	.2	.4	.3

We have two results we can use:

Checking (4.10). To check irreducibility we note that starting from 0, 1, or 5 we can get to 2, 3, 4 and 5 in one step and to 0 and 1 in two steps by going through 1 or 2. From 2 or 3 we can get to 0, 1, and 2 in one step and to 3, 4, and 5 in two steps by going through 0. Finally from 4 we can get to 1, 2, 3, and 4 in one step and to 0 or 5 in two steps by going through 2 or 0 respectively. To check aperiodicity, we note that $p(5, 5) > 0$ so 5 is aperiodic.

Checking (4.11). With a calculator we can compute p^2 :

	0	1	2	3	4	5
0	.05	.12	.22	.28	.24	.09
1	.05	.12	.22	.28	.24	.09
2	.09	.12	.16	.14	.28	.21
3	.15	.22	.27	.15	.12	.09
4	.10	.19	.29	.26	.13	.03
5	.05	.12	.22	.28	.24	.09

All entries are positive so (4.11) applies.

Doubly Stochastic Chains. A Markov chain is defined by the condition that $\sum_j p(i, j) = 1$. Suppose that in addition, we have $\sum_i p(i, j) = 1$. If the chain has N states then the stationary distribution is $\pi(i) = 1/N$ since

$$\sum_i \pi(i)p(i, j) = \frac{1}{N} \sum_i p(i, j) = \frac{1}{N}$$

To illustrate this consider

Example 4.18. Tiny Board Game. Consider a circular board game with only six spaces $\{0, 1, 2, 3, 4, 5\}$. On each turn we decide how far to move by flipping two coins and then moving one space for each heads. Here we consider 5 to be adjacent to 0, so if we are there and get two heads then the result is $5 + 2 \bmod 6 = 1$, where $i + k \bmod 6$ is the remainder when $i + k$ is divided by 6. In this case the transition probability is

	0	1	2	3	4	5
0	1/4	1/2	1/4	0	0	0
1	0	1/4	1/2	1/4	0	0
2	0	0	1/4	1/2	1/4	0
3	0	0	0	1/4	1/2	1/4
4	1/4	0	0	0	1/4	1/2
5	1/2	1/4	0	0	0	1/4

It is clear that the columns add to one, so the stationary distribution is uniform. To check the hypothesis of the convergence theorem, we note that after 3 turns we will have moved between 0 and 6 spaces so $p^3(i, j) > 0$.

Example 4.19. Mathematician's Monopoly. The game Monopoly is played on a game board that has 40 spaces arranged around the outside of a square. The squares have names like *Reading Railroad* and *Park Place* but we will number the squares 0 (*Go*), 1 (*Baltic Avenue*), ... 39 (*Boardwalk*). In Monopoly you roll two dice and move forward a number of spaces equal to the sum. For the moment, we will ignore things like *Go to Jail*, *Chance*, and other squares that make the transitions complicated and formulate the dynamics as following. Let

r_k be the probability that the sum of two dice is k ($r_2 = 1/36$, $r_3 = 2/36$, \dots , $r_7 = 6/36$, \dots , $r_{12} = 1/36$) and let

$$p(i, j) = r_k \quad \text{if } j = i + k \pmod{40}$$

where $i + k \pmod{40}$ is the remainder when $i + k$ is divided by 40. To explain suppose that we are sitting on *Park Place* $i = 37$ and roll $k = 6$. $37 + 6 = 43$ but when we divide by 40 the remainder is 3, so $p(37, 3) = r_6 = 5/36$.

This example is larger but has the same structure as the previous example. Each row has the same entries but shift one unit to the right each time with the number that goes off the right edge emerging in the 0 column. This structure implies that each entry in the row appears once in each column and hence the sum of the entries in the column is 1, and the stationary distribution is uniform. To check the hypothesis of the convergence theorem note that in four rolls you can move forward by 8 to 48 squares, so $p^4(i, j) > 0$ for all i and j .

The real game of Monopoly has two complications:

- Square 30 is “Go to Jail,” which sends you to square 10. You can buy your way out of jail but in the results we report below, we assume that you are cheap. If you roll a double then you get out for free. If you don’t get doubles in three tries you have to pay.
- There are three *Chance* squares at 7, 12, and 36 (diamonds on the graph), and three *Community Chest* squares at 2, 17, 33 (squares on the graph), where you draw a card, which can send you to another square.

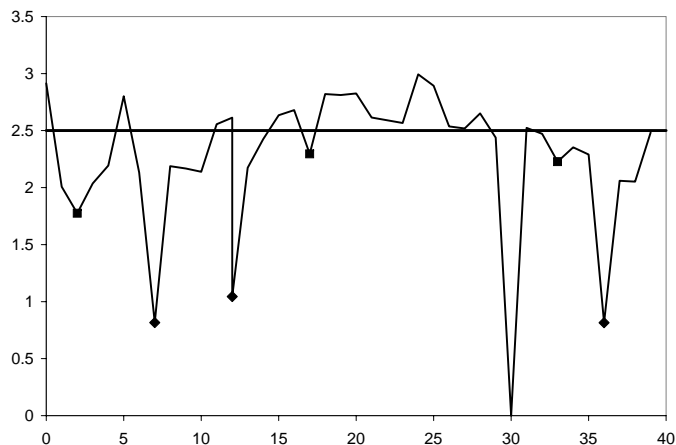


Figure 4.1: Stationary distribution for Monopoly squares, except “In Jail”.

The graph in Figure 4.1 gives the long run frequencies of being in different squares on the Monopoly board at the end of your turn, as computed by simulation. To make things easier to see we have removed the 9.46% chance of being

In Jail to make the probabilities easier to see. The value reported for 10 is the 2.14% probability of *Just Visiting Jail*, i.e., being brought there by the roll of the dice. Square 30, *Go to Jail*, has probability 0 for the obvious reasons. The other three lowest values occur for *Chance* squares. Due to the transition from 30 to 10, frequencies for squares near 20 are increased relative to the average of 2.5% while those after 30 or before 10 are decreased. Squares 0 (*Go*) and 5 (*Reading Railroad*) are exceptions to this trend since there are Chance card that instruct you to go there.

This example is not clear.

Example 4.20. Landing on Go. What is the long run probability that during one trip around the board in Mathematician's Monopoly that we land on Go?

Intuitive Solution. Consider instead an infinite board with squares $0, 1, 2, \dots$. Rolling two dice we move an average of 7 square per turn, so in the long run we visit $1/7$ of the squares.

Formal Solution. Let $X_n = 0$ if we visit square n on the infinite board. For the squares m we do not visit, we define X_m to be the number of steps to the next square we hit. For example if the first three rolls were 4, 6, and 3.

$$\begin{array}{cccccccccccccccc} X_n & 0 & 3 & 2 & 1 & 0 & 5 & 4 & 3 & 2 & 1 & 0 & 2 & 1 & 0 \\ n & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \end{array}$$

Let q_k be the probability we move forward k squares on one roll and suppose $q_0 = 0$. Then X_n is a Markov chain with

$$\begin{aligned} p_{j,j-1} &= 1 & j > 0 \\ p_{0,k-1} &= q_k \end{aligned}$$

To check the second note that in the example the first roll is 4 but $X_1 = 3$. Let $r_k = \sum_{j=k+1}^{\infty} q_j$. We claim that $rp = r$. To check this, note that to end at $k \geq 0$ there are two possibilities: come down from $k+1$ or jump up from 0, so

$$(rp)_k = r_{k+1} \cdot 1 + r_0 \cdot q_{k+1} = r_k$$

since $r_0 = \sum_{j=1}^{\infty} q_j = 1$ and $q_{k+1} + \sum_{j=k+2}^{\infty} q_j = r_k$.

To turn r into a stationary distribution we need to divide by $\sum_{k=0}^{\infty} r_k$. To evaluate the sum we write

$$\sum_{k=0}^{\infty} \sum_{j=k+1}^{\infty} q_j = \sum_{j=0}^{\infty} \sum_{k=0}^{j-1} q_j = \sum_{j=0}^{\infty} j q_j = \mu$$

Thus the stationary distribution is $\pi_k = r_k / \mu$.

When q_k is the probability the sum of two dice is k , then X_n is irreducible on the state space $\{0, 1, \dots, 11\}$. $p(0, 0) = q_1 > 0$ so X_n is a periodic and we can conclude $P(X_n = 0) \rightarrow 1/\mu = 1/7$.

4.5 Gambler's Ruin

As we have said earlier, the long run behavior of the Gambler's Ruin chain and the Wright-Fisher model with no mutation are not very exciting. After a while the chain enters one of the absorbing states (0 and N) and stays there forever. Our first question is what is the probability that the chain gets absorbed at N before hitting 0, i.e., what is the probability that the gambler reaches his goal?

Example 4.21. Flipping coins. A gambler is betting \$1 each time on the outcome of the flip of a fair coin. He has \$10 and will stop when he has \$25. What is the probability he will reach his goal before he runs out of money?

Even though we are only interested in what happens when we start at 10, to solve this problem we must compute $h(x)$ = the probability of reaching N before 0 starting from x . Of course $h(0) = 0$ and $h(N) = 1$. For $0 < x < N$, considering what happens on the first step gives

$$h(x) = \frac{1}{2}h(x-1) + \frac{1}{2}h(x+1)$$

In words $h(x)$ is the average of $h(x-1)$ and $h(x+1)$. Multiplying by 2, moving $h(x-1)$ to the left, and one of the $h(x)$'s to the right we have

$$h(x) - h(x-1) = h(x+1) - h(x)$$

This says that the slope of h is constant or the graph of h is a straight line. Since $h(0) = 0$ and $h(N) = 1$, the slope must be $1/N$ and

$$h(x) = x/N \tag{4.12}$$

Thus the answer to our question is $10/25$.

To see that this is reasonable, note that since we are playing a fair game (i.e., the average winnings on any play is 0) the average amount of money we have at any time is the \$10 we started with. When the game ends we will have \$25 with probability p and \$0 with probability $1-p$. For the expected value to be \$10 we must have $p = 10/25$. This calculation extends easily to the general case: when the game ends we will have \$ N with probability p and \$0 with probability $1-p$. If we start with \$ x then the expected value at the end should also be \$ x and we must have $p = x/N$.

Example 4.22. Wright-Fisher model. As described in Example 4.2, if we let X_n be the number of A alleles at time n , then X_n is a Markov chain with transition probability

$$p(x, y) = \binom{N}{y} \left(\frac{x}{N}\right)^j \left(1 - \frac{x}{N}\right)^{N-y}$$

0 and N are absorbing states. What is the probability the chain ends up in N starting from x .

Extending the reasoning in the previous example, we see that if $h(x)$ is the probability of getting absorbed in state N when we start in state x then $h(0) = 0$, $h(N) = 1$, and for $0 < x < N$

$$h(x) = \sum_y p(x, y)h(y) \quad (4.13)$$

In words if we jump from x to y on the first step then our absorption probability becomes $h(y)$.

In the Wright-Fisher model $p(x, y)$ is the binomial distribution for N independent trials with success probability x/N , so the expected number of A 's after the transition is x , i.e., the expected number of A 's remains constant in time. Using the reasoning from the previous example, we guess

$$h(y) = y/N$$

Clearly, $h(0) = 0$ and $h(N) = 1$. To check (4.13) we note that

$$\sum_y p(x, y)y/N = x/N$$

since the mean of the binomial is x .

The formula $h(y) = y/N$ says that if we start with y A 's in the population then the probability we will end with a population of all A 's (an event called "fixation" in genetics) is y/N , the fraction of the population that is A . The case $y = 1$ is a famous result due to Kimura: the probability of fixation of a new mutation is $1/N$. If we suppose that each individual experiences mutations at rate μ , then since there are N individuals, new mutations occur at a total rate $N\mu$. Since each mutation achieves fixation with probability $1/N$, the rate at which mutations become fixed is μ independent of the size of population.

Example 4.23. Roulette. Suppose now that the gambler is playing roulette where he will win \$1 with probability $p = 18/38$ and lose \$1 with probability $1 - p = 20/38$ each time. He has \$25 and will stop when he has \$50. What is the probability he will reach his goal before he runs out of money? Note if this was a fair game his success probability would be 0.5.

Again we let $h(x)$ = the probability of reaching N before 0 starting from x . Of course $h(0) = 0$ and $h(N) = 1$. For $0 < x < N$, (4.13) implies

$$h(x) = (1 - p)h(x - 1) + ph(x + 1)$$

Moving $(1 - p)h(x - 1)$ to the left, and $ph(x)$ to the right we have

$$(1 - p)(h(x) - h(x - 1)) = p(h(x + 1) - h(x))$$

which rearranges to

$$h(x + 1) - h(x) = \frac{1 - p}{p}(h(x) - h(x - 1)) \quad (\star)$$

We know that $h(0) = 0$. We don't know $h(1)$, so we let $h(1) = c$. Using (\star) repeatedly we have

$$\begin{aligned} h(2) - h(1) &= \frac{1-p}{p}(h(1) - h(0)) = \left(\frac{1-p}{p}\right) c \\ h(3) - h(2) &= \frac{1-p}{p}(h(2) - h(1)) = \left(\frac{1-p}{p}\right)^2 c \\ h(4) - h(3) &= \frac{1-p}{p}(h(3) - h(2)) = \left(\frac{1-p}{p}\right)^3 c \end{aligned}$$

From this it should be clear that

$$h(x+1) - h(x) = \left(\frac{1-p}{p}\right)^x c$$

Writing $r = (1-p)/p$ to simplify, and recalling $h(0) = 0$ we have

$$\begin{aligned} h(y) &= h(y) - h(0) = \sum_{x=1}^y h(x) - h(x-1) \\ &= (r^{y-1} + r^{y-2} + \dots + r + 1)c = \frac{r^y - 1}{r - 1} \cdot c \end{aligned}$$

since $(r^{y-1} + r^{y-2} + \dots + r + 1)(r - 1) = r^y - 1$. We want $h(N) = 1$ so we must have $c = (r - 1)/(r^N - 1)$. It follows that

$$h(y) = \frac{r^y - 1}{r^N - 1} = \frac{\left(\frac{1-p}{p}\right)^y - 1}{\left(\frac{1-p}{p}\right)^N - 1} \quad (4.14)$$

To see what this says for our roulette example we take $p = 18/38$, $x = 25$, $N = 50$. In this case $(1-p)/p = 10/9$ so the probability we succeed is

$$\frac{(10/9)^{25} - 1}{(10/9)^{50} - 1} = \frac{12.929}{193.03} = 0.067$$

compared to 0.5 for the fair game.

Now let's turn things around and look at the game from the viewpoint of the casino, i.e., $p = 20/38$. Suppose that the casino starts with the rather modest capital of $x = 100$. (4.14) implies that the probability they will reach N before going bankrupt is

$$\frac{(9/10)^{100} - 1}{(9/10)^N - 1}$$

If we let $N \rightarrow \infty$, $(9/10)^N \rightarrow 0$ so the answer converges to

$$1 - (9/10)^{100} = 1 - 2.656 \times 10^{-5}$$

If we increase the capital to \$200 then the failure probability is squared, since to become bankrupt we must first lose \$100 and then lose our second \$100. In this case the failure probability is incredibly small: $(2.656 \times 10^{-5})^2 = 7.055 \times 10^{-10}$.

4.6 Absorbing Chains

In this section we consider general Markov chains with absorbing states. The two questions of interest are: "where does the chain get absorbed?" and "how long does it take to get there?" We begin with a simple example.

Example 4.24. Two year college. At a local two year college. 60% of freshmen become sophomores, 25% remain freshmen, and 15% drop out. 70% of sophomores graduate and transfer to a four year college, 20% remain sophomores and 10% drop out. What fraction of new students eventually graduate?

We use a Markov chain with state space $1 = \text{freshman}$, $2 = \text{sophomore}$, $G = \text{graduate}$, $D = \text{dropout}$. The transition probability is

	1	2	G	D
1	0.25	0.6	0	0.15
2	0	0.2	0.7	0.1
G	0	0	1	0
D	0	0	0	1

Let $h(x)$ be the probability that a student currently in state x eventually graduates. By considering what happens on one step

$$\begin{aligned} h(1) &= 0.25h(1) + 0.6h(2) \\ h(2) &= 0.2h(2) + 0.7 \end{aligned}$$

since $h(G)=1$ and $h(D)=0$

so $h(2) = 0.7/0.8 = 0.875$ and $h(1) = (0.6)/(0.75)h(2) = 0.7$.

To check these answers we will look at powers of the transition probability

	1	2	G	D
1	0.0625	0.27	0.42	0.2475
$p^2 = 2$	0	0.04	0.84	0.12
G	0	0	1	0
D	0	0	0	1

From this we see that the fraction of freshmen who graduate in two years is $p^2(1, G) = 0.42 = 0.6(0.7) = p(1, 2)p(2, G)$. After three years

	1	2	G	D
1	0.015625	0.0915	0.609	0.283875
$p^3 = 2$	0	0.008	0.868	0.124
G	0	0	1	0
D	0	0	0	1

while after 6 years very few people are still trying to get their degrees

	1	2	G	D
1	0.000244	0.002161	0.697937	0.283875
$p^6 = 2$	0	0.000064	0.874944	0.124492
G	0	0	1	0
D	0	0	0	1

Example 4.25. Tennis. In tennis the winner of a game is the first player to win four points, unless the score is 4–3, in which case the game must continue until one player wins by two points. Suppose that the game has reached the point where one player is trying to get two points ahead to win and that the server will independently win the point with probability 0.6. What is the probability the server will win the game if the score is tied 3–3? if she is ahead by one point? Behind by one point?

We formulate the game as a Markov chain in which the state is the difference of the scores. The state space is 2, 1, 0, –1, –2 with 2 (win for server) and –2 (win for opponent). The transition probability is

$$p = \begin{array}{c} \begin{array}{ccccc} & 2 & 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & 0 & 0 & 0 \\ 1 & .6 & 0 & .4 & 0 & 0 \\ 0 & 0 & .6 & 0 & .4 & 0 \\ -1 & 0 & 0 & .6 & 0 & .4 \\ -2 & 0 & 0 & 0 & 0 & 1 \end{array} \end{array}$$

If we let $h(x)$ be the probability of the server winning when the score is x then

$$h(x) = \sum_y p(x, y)h(y)$$

with $h(2) = 1$ and $h(-2) = 0$. This involves solving three equations in three unknowns. The computations become much simpler if we look at

$$p^2 = \begin{array}{c} \begin{array}{ccccc} & 2 & 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & 0 & 0 & 0 \\ 1 & .6 & .24 & 0 & .16 & 0 \\ 0 & .36 & 0 & .48 & 0 & .16 \\ -1 & 0 & .36 & 0 & .24 & .4 \\ -2 & 0 & 0 & 0 & 0 & 1 \end{array} \end{array}$$

From p^2 we see that

$$h(0) = 0.36 + 0.48h(0) \quad \text{since } h(2)=1 \text{ and } h(-2)=0$$

so $h(0) = 0.36/0.52 = 0.6923$. By considering the outcome of the first point we see that $h(1) = 0.6 + 0.4h(0) = 0.8769$ and $h(-1) = 0.6h(0) = 0.4154$.

General solution. Suppose that the server wins each point with probability w . If the game is tied then after two points, the server will have won with probability w^2 , lost with probability $(1-w)^2$, and returned to a tied game with probability $2w(1-w)$, so $h(0) = w^2 + 2w(1-w)h(0)$. Since $1 - 2w(1-w) = w^2 + (1-w)^2$, solving gives

$$h(0) = \frac{w^2}{w^2 + (1-w)^2}$$

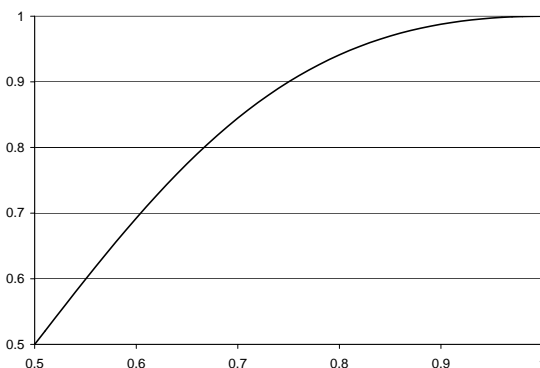


Figure 4.2: Probability of winning a tied game as a function of the probability of winning a point.

Absorption times

Example 4.26. An office computer is in one of three states, working (W), being repaired (R), or scrapped (S). If the computer is working one day the probability it will be working the next day is .995 and the probability it will need repair is .005. If it is being repaired then probability it is working the next day is .9, the probability it still needs repair the next day is .05 and the probability it will be scrapped is .05. What is the average number of working days until a computer is scrapped?

Leaving out the absorbing state of being scrapped the reduced transition probability is

$$r = \begin{array}{cc} & \begin{array}{c} W \\ R \end{array} \\ \begin{array}{c} W \\ R \end{array} & \begin{array}{cc} .995 & .005 \\ .90 & .05 \end{array} \end{array}$$

To see that this is enough of the matrix to compute what we want note that $r^n(W, W)$ gives the probability a computer is working on day n , so $\sum_{n=0}^{\infty} r^n(W, W)$ gives the expected number of days that it is working. To see this, let $Y_n = 1$ if the computer is working on day n and 0 otherwise, then $Y = \sum_{n=0}^{\infty} Y_n$ is the number of days the computer is working

$$EY = E \sum_{n=0}^{\infty} Y_n = \sum_{n=0}^{\infty} EY_n = \sum_{n=0}^{\infty} r^n(W, W)$$

By analogy with the geometric series $\sum_{n=0}^{\infty} x^n = 1/(1-x)$ we can guess that

$$\sum_{n=0}^{\infty} r^n = (I - r)^{-1} \quad (4.15)$$

where I is the identity matrix, and $r^0 = I$. To check this we note that

$$(I - r) \sum_{n=0}^{\infty} r^n = \sum_{n=0}^{\infty} r^n - \sum_{n=1}^{\infty} r^n = r^0 = I$$

Computing the inverse

$$(I - r)^{-1} = \begin{array}{cc} & \begin{array}{cc} W & R \end{array} \\ \begin{array}{c} W \\ R \end{array} & \begin{array}{cc} 3800 & 20 \\ 3600 & 20 \end{array} \end{array}$$

we see that on the average a working computer will work for 3800 days and will spend 20 days being repaired.

$(I - r)^{-1}$ is sometimes called the **fundamental matrix** because it is the key to computing many quantities for absorbing Markov chains.

Example 4.27. A local cable company classifies their customers according to how many months overdue their bill is: 0,1,2,3. Accounts that are three months overdue are discontinued (D) if they are not paid. The company estimates that transitions occur according to the following probabilities:

	0	1	2	3	D
0	.9	.1	0	0	0
1	.8	0	.2	0	0
2	.7	0	0	.3	0
3	.6	0	0	0	.4
D	0	0	0	0	1

What is the expected number of months for a new customer (i.e., one who starts in state 0) to have their service discontinued?

Let r be the reduced 4×4 matrix of transitions between the non-absorbing states, 0-3.

$$(I - r)^{-1} = \begin{array}{cccc} 416.66 & 41.66 & 8.33 & 2.5 \\ 406.66 & 41.66 & 8.33 & 2.5 \\ 366.66 & 36.66 & 8.33 & 2.5 \\ 250 & 25 & 5 & 2.5 \end{array}$$

The first row gives **the expected number of visits** to each of the four states starting from 0 so the expected time is $416.66 + 41.66 + 8.33 + 2.5 = 469.16$.

Returning to our first two examples:

Example 4.28. Two year college. Consider the transition probability in Example 4.24. How many years on the average does it take for a freshman to graduate or dropout?

Removing the absorbing states from the transition probability, we obtain the reduced matrix

$$r = \begin{array}{cc|cc} & & 1 & 2 \\ \hline 1 & 0.25 & 0.6 & \\ 2 & 0 & 0.2 & \end{array}$$

From this we compute

$$(I - r)^{-1} = \begin{array}{cc|cc} & & 1 & 2 \\ \hline 1 & 1.33 & 1 & \\ 2 & 0 & 1.25 & \end{array}$$

so, on the average, a freshman takes $1.33 + 1 = 2.33$ years to either graduate or drop out.

Example 4.29. Tennis. Consider the transition probability in Example 4.25. Suppose that game is tied 3-3? How many more points do we expect to see before the game ends?

Removing the absorbing states from the transition probability, we obtain the reduced matrix

$$r = \begin{array}{ccc|ccc} & & & 1 & 0 & -1 \\ \hline & 1 & 0 & .4 & 0 & \\ 0 & .6 & 0 & .4 & & \\ -1 & 0 & .6 & 0 & & \end{array}$$

From this we compute

$$(I - r)^{-1} = \begin{array}{ccc|ccc} & & & 1 & 0 & -1 \\ \hline 1 & 1.4615 & .7692 & .307 & & \\ 0 & 1.1538 & 1.09230 & .769 & & \\ -1 & .6923 & 1.1538 & 1.461 & & \end{array}$$

so, on the average, a tied game requires $1.1538 + 1.09230 + .769 = 3.8458$ points to be completed.

The fundamental matrix can also be used to compute the probability of winning the game. To see this we note that in order to end in state 2, the chain must wander among the nonabsorbing states for some number of times n and then jump from some state y to state 2, i.e.,

$$h(x) = \sum_y \sum_{n=0}^{\infty} r^n(x, y)p(y, 2) = (I - r)^{-1}(x, y)p(y, 2)$$

In the case of the tennis chain. $p(y, 2) = 0$ unless $y = 1$ and in this case $p(1, 2) = 0.6$ so

$$h(x) = 0.6(I - p)^{-1}(x, 1)$$

Multiplying the first column of the previous matrix by 0.6 we get the answers we found in Example 4.25:

$$.8769 \quad .6923 \quad .4154$$

4.7 Exercises

Transition Probabilities

1. What values of x, y, z will make these matrices transition probabilities:

$$\begin{array}{ccc} & \begin{matrix} .5 & .1 & x \end{matrix} & \\ \text{a. } & \begin{matrix} y & .2 & .4 \\ .3 & z & .1 \end{matrix} & \end{array} \quad \begin{array}{ccc} & \begin{matrix} x & .1 & .7 \end{matrix} & \\ \text{b. } & \begin{matrix} .2 & .3 & y \\ .6 & z & .2 \end{matrix} & \end{array}$$

2. A red urn contains 2 red marbles and 3 blue marbles. A blue urn contains 1 red marble and 4 blue marbles. A marble is selected from an urn, the marble is returned to the urn from which it was drawn and the next marble is drawn from the urn with the color that was drawn. (a) Write the transition probability for this chain. (b) Suppose the first marble is drawn from the red urn. What is the probability the third one will be drawn from the blue urn?

3. At Llenroc College, 63% of freshmen who are pre-med switch to a liberal arts major, while 18% of liberal arts majors switch to being pre-med. If the incoming freshman class is 60% pre-med and 40% liberal arts majors, what fraction graduate as pre-med?

4. A person is flipping a coin repeatedly. Let X_n be the outcome of the two previous coin flips at time n , for example the state might be HT to indicate that the last flip was T and the one before that was H . (a) compute the transition probability for the chain. (b) Find p^2 .

5. A taxicab driver moves between the airport A and two hotels B and C according to the following rules. If he is at the airport, he will go to one of the two hotels next with equal probability. If at a hotel then he returns to the airport with probability $3/4$ and goes to the other hotel with probability $1/4$. (a) Find the transition matrix for the chain. (b) Suppose the driver begins at the airport at time 0. Find the probability for each of his three possible locations at time 2 and the probability he is at hotel B at time 3.

6. Consider a gambler's ruin chain with $N = 4$. That is, if $1 \leq i \leq 3$, $p(i, i+1) = 0.4$, and $p(i, i-1) = 0.6$, but the endpoints are absorbing states: $p(0, 0) = 1$ and $p(4, 4) = 1$ Compute $p^3(1, 4)$ and $p^3(1, 0)$.

7. An outdoor restaurant in a resort town closes when it rains. From past records it was found that from May to September, when it rains one day the probability that it rains the next is 0.4; when it does not rain one day it rains the next with probability 0.1. (a) Write the transition matrix. (b) If it rained on Thursday what is the probability that it will rain on Saturday? on Sunday?

8. Market research suggests that in a five year period 8% of people with cable television will get rid of it, and 26% of those without it will sign up for it. Compare the predictions of the Markov chain model with the following data on the fraction of people with cable TV: 56.4% in 1990, 63.4% in 1995, and 68.0% in 2000.

9. A sociology professor postulates that in each decade 8% of women in the work force leave it and 20% of the women not in it begin to work. Compare the predictions of his model with the following data on the percentage of women working: 43.3% in 1970, 51.5% in 1980, 57.5% in 1990, and 59.8% in 2000.
10. The following transition probability describes the migration patterns of birds between three habitats

	1	2	3
1	.75	.15	.10
2	.07	.85	.08
3	.05	.15	.80

If there are 1000 birds in each habitat at the beginning of the first year, how many do we expect to be in each habitat at the end of the year? at the end of the second year?

Convergence to Equilibrium

11. A car rental company has rental offices at both Kennedy and LaGuardia airports. Assume that a car rented at one airport must be returned to one of the two airports. If the car was rented at LaGuardia the probability it will be returned there is 0.8; for Kennedy the probability is 0.7. Suppose that we start with $1/2$ of the cars at each airport and that each week all of the cars are rented once. (a) What is the fraction of cars at LaGuardia airport at the end of the first week? (b) at the end of the second? (c) in the long run?
12. The 1990 census showed that 36% of the households in the District of Columbia were homeowners while the remainder were renters. During the next decade 6% of the homeowners became renters and 12% of the renters became homeowners. (a) What percentage were homeowners in 2000? in 2010? (b) If these trends continue what will be the long run fraction of homeowners?
13. Most railroad cars are owned by individual railroad companies. When a car leaves its home railroad's trackage, it becomes part of the national pool of cars and can be used by other railroads. A particular railroad found that each month 15% of its boxcars on its home trackage left to join the national pool and 40% of its cars in the national pool were returned to its home trackage. A company begins on January 1 with all of its cars on its home trackage. What fraction will be there on March 1? At the end of the year? In the long run what fraction of a company's cars will be on its home trackage.
14. A rapid transit system has just started operating. In the first month of operation, it was found that 25% of commuters are using the system while 75% are travelling by automobile. Suppose that each month 10% of transit users go back to using their cars, while 30% of automobile users switch to the transit system. (a) Compute the three step transition probability p^3 . (b) What will be the fractions using rapid transit in the fourth month? (c) In the long run?
15. A regional health study indicates that from one year to the next, 75% percent of smokers will continue to smoke while 25% will quit. 8% of those

who stopped smoking will resume smoking while 92% will not. If 70% of the population were smokers in 1995, what fraction will be smokers in 1998? in 2005? in the long run?

16. The town of Mythica has a “free bikes for the people program.” You can pick up bikes at the library (L), the coffee shop (C) or the cooperative grocery store (G). The director of the program has determined that if a bike is picked up at the library ends up at the coffee shop with probability 0.2 and at the grocery store with probability 0.3. A bike from the coffee shop will go to the library with probability 0.4 and to the grocery store with probability 0.1. A bike from the grocery store will go to the library or the coffee shop with probability 0.25 each. On Sunday there are an equal number of bikes at each place. (a) What fraction of the bikes are at the three locations on Tuesday? (b) on the next Sunday? (c) In the long run what fraction are at the three locations?

17. Bob eats lunch at the campus food court every week day. He either eats Chinese food, Quesadila, or Salad. His transition matrix is

	<i>C</i>	<i>Q</i>	<i>S</i>
<i>C</i>	.15	.6	.25
<i>Q</i>	.4	.1	.5
<i>S</i>	.1	.3	.6

He had Chinese food on Monday. What are the probabilities for his three meal choices on Friday (four days later).

Asymptotic Behavior: Two state chains

18. Census results reveal that in the United States 80% of the daughters of working women work and that 30% of the daughters of nonworking women work. (a) Write the transition probability for this model. (b) In the long run what fraction of women will be working?

19. Three of every four trucks on the road are followed by a car, while only one of every five cars is followed by a truck. What fraction of vehicles on the road are trucks?

20. In a test paper the questions are arranged so that $3/4$'s of the time a True answer is followed by a True, while $2/3$'s of the time a False answer is followed by a False. You are confronted with a 100 question test paper. Approximately what fraction of the answers will be True.

21. When a basketball player makes a shot then he tries a harder shot the next time and hits (H) with probability 0.4, misses (M) with probability 0.6. When he misses he is more conservative the next time and hits (H) with probability 0.7, misses (M) with probability 0.3. (a) Write the transition probability for the two state Markov chain with state space $\{H, M\}$. (b) Find the long-run fraction of time he hits a shot.

22. A regional health study shows that from one year to the next 76% of the people who smoked will continue to smoke and 24% will quite. 8% of those

who do not smoke will start smoking while 92% of those who do not smoke will continue to be nonsmokers. In the long run what fraction of people will be smokers?

23. In unprofitable times corporations sometimes suspend dividend payments. Suppose that after a dividend has been paid the next one will be paid with probability 0.9, while after a dividend is suspended the next one will be suspended with probability 0.6. In the long run what is the fraction of dividends that will be paid?

24. A university computer room has 30 terminals. Each day there is a 3% chance that a given terminal will break and a 72% chance that a given broken terminal will be repaired. Assume that the fates of the various terminals are independent. In the long run what is the distribution of the number of terminals that are broken.

Asymptotic Behavior: Three or more states

25. A plant species has red, pink, or white flowers according to the genotypes RR, RW, and WW, respectively. If each of these genotypes is crossed with a pink (RW) plant then the offspring fractions are

	RR	RW	WW
RR	.5	.5	0
RW	.25	.5	.25
WW	0	.25	.5

What is the long run fraction of plants of the three types?

26. A certain town never has two sunny days in a row. Each day is classified as rainy, cloudy, or sunny. If it is sunny one day then it is equally likely to be cloudy or rainy the next. If it is cloudy or rainy, then it remains the same 1/2 of the time, but if it changes it will go to either of the other possibilities with probability 1/4 each. In the long run what proportion of days in this town are sunny? cloudy? rainy?

27. A midwestern university has three types of health plans: a health maintenance organization (HMO), a preferred provider organization (PPO), and a traditional fee for service plan (FFS). In 2000, the percentages for the three plans were HMO:30%, PPO:25%, and FFS:45%. Experience dictates that people change plans according to the following transition matrix

	HMO	PPO	FFS
HMO	.85	.1	.05
PPO	.2	.7	.1
FFS	.1	.3	.6

(a) What will be the percentages for the three plans in 2001? (b) What is the long run fraction choosing each of the three plans?

28. A sociologist studying living patterns in a certain region determines that the pattern of movement between urban (U), suburban (S), and rural areas (R) is given by the following transition matrix.

	<i>U</i>	<i>S</i>	<i>R</i>
<i>U</i>	0.86	0.08	0.06
<i>S</i>	0.05	0.88	0.07
<i>R</i>	0.03	0.05	0.92

In the long run what fraction of the population will live in the three areas.

29. In a large metropolitan area, commuters either drive alone (A), carpool (C), or take public transportation (T). A study showed that 80% of those who drive alone will continue to do so next year, while 15% will switch to carpooling and 5% will use public transportation. 90% of those who carpool will continue, while 5% will drive alone and 5% will use public transportation. 85% of those who use public transportation will continue, while 10% will carpool, and 5% will drive alone. Write the transition probability for the model. In the long run what fraction of commuters will use the three types of transportation?

30. In a particular county voters declare themselves as members of the Republican, Democrat, or Green party. No voters change directly from the Republican to Green party or vice versa. In a given year 15% of Republicans and 5% of Green party members will become Democrats, while 5% of Democrats switch to the Republican party and 10% to the Green party. Write the transition probability for the model. In the long run what fraction of voters will belong to the three parties.

31. (a) Three telephone companies *A*, *B*, and *C* compete for customers. Each year *A* loses 5% of its customers to *B* and 20% to *C*; *B* loses 15% of its customers to *A* and 20% to *C*; *C* loses 5% its customers to *A* and 10% to *B*. (a) Write the transition matrix for the model. (b) What is the limiting market share for each of these companies?

32. An auto insurance company classifies its customers in three categories: poor, satisfactory and preferred. No one moves from poor to preferred or from preferred to poor in one year. 40% of the customers in the poor category become satisfactory, 30% of those in the satisfactory category moves to preferred, while 10% become poor; 20% of those in the preferred category are downgraded to satisfactory. (a) Write the transition matrix for the model. (b) What is the limiting fraction of drivers in each of these categories?

33. A professor has two light bulbs in his garage. When both are burned out, they are replaced, and the next day starts with two working light bulbs. Suppose that when both are working, one of the two will go out with probability .02 (each has probability .01 and we ignore the possibility of losing two on the same day). However, when only one is there, it will burn out with probability .05. What is the long-run fraction of time that there is exactly one bulb working?

34. An individual has three umbrellas, some at her office, and some at home. If she is leaving home in the morning (or leaving work at night) and it is raining, she will take an umbrella, if one is there. Otherwise, she gets wet. Assume that independent of the past, it rains on each trip with probability 0.2. To formulate a Markov chain, let X_n be the number of umbrellas at her current location. (a) Find the transition probability for this Markov chain. (b) Calculate the limiting fraction of time she gets wet.

35. At the end of a month, a large retail store classifies each of its customer's accounts according to current (0), 30–60 days overdue (1), 60–90 days overdue (2), more than 90 days overdue (3). Their experience indicates that the accounts move from state to state according to a Markov chain with transition probability matrix:

	0	1	2	3
0	.9	.1	0	0
1	.8	0	.2	0
2	.5	0	0	.5
3	.1	0	0	.9

In the long run what fraction of the accounts are in each category?

36. At the beginning of each day, a piece of equipment is inspected to determine its working condition, which is classified as state 1 = new, 2, 3, or 4 = broken. We assume the state is a Markov chain with the following transition matrix:

	1	2	3	4
1	.95	.05	0	0
2	0	.9	.1	0
3	0	0	.875	.125

(a) Suppose that a broken machine requires three days to fix. To incorporate this into the Markov chain we add states 5 and 6 and suppose that $p(4,5) = 1$, $p(5,6) = 1$, and $p(6,1) = 1$. Find the fraction of time that the machine is working. (b) Suppose now that we have the option of performing preventative maintenance when the machine is in state 3, and that this maintenance takes one day and returns the machine to state 1. This changes the transition probability to

	1	2	3
1	.95	.05	0
2	0	.9	.1
3	1	0	0

Find the fraction of time the machine is working under this new policy.

37. To make a crude model of a forest we might introduce states 0 = grass, 1 = bushes, 2 = small trees, 3 = large trees, and write down a transition matrix

like the following:

	0	1	2	3
0	1/2	1/2	0	0
1	1/24	7/8	1/12	0
2	1/36	0	8/9	1/12
3	1/8	0	0	7/8

The idea behind this matrix is that if left undisturbed a grassy area will see bushes grow, then small trees, which of course grow into large trees. However, disturbances such as tree falls or fires can reset the system to state 0. Find the limiting fraction of land in each of the states.

38. Five white balls and five black balls are distributed in two urns in such a way that each urn contains five balls. At each step we draw one ball from each urn and exchange them. Let X_n be the number of white balls in the left urn at time n . (a) Compute the transition probability for X_n . (b) Find the stationary distribution and show that it corresponds to picking five balls at random to be in the left urn.

Absorbing Chains

39. Two competing companies are trying to buy up all of the farms in a certain area to build houses. In each year 10% of farmers sell to company 1, 20% sell to company 2, and 70% keep farming. Neither company ever sells any of the farms that they own. Eventually all of the farms will be sold. How many will be owned by company 1?

40. A warehouse has a capacity to hold four items. If the warehouse is neither full nor empty, the number of items in the warehouse changes whenever a new item is produced or an item is sold. Suppose that (no matter when we look) the probability that the next event is “a new item is produced” is $2/3$ and that the new event is a “sale” is $1/3$. If there is currently one item in the warehouse, what is the probability that the warehouse will become full before it becomes empty.

41. The Microsoft company gives each of its employees the title of programmer (P) or project manager (M). In any given year 70% of programmers remain in that position 20% are promoted to project manager and 10% are fired (state X). 95% of project managers remain in that position while 5% are fired. How long on the average does a programmer work before they are fired?

42. At a nationwide travel agency, newly hired employees are classified as beginners (B). Every six months the performance of each agent is reviewed. Past records indicate that transitions through the ranks to intermediate (I) and qualified (Q) are according to the following Markov chain, where F indicates workers that were fired:

	B	I	Q	F
B	.45	.4	0	.15
I	0	.6	.3	.1
Q	0	0	1	0
F	0	0	0	1

(a) What fraction are eventually promoted? (b) What is the expected time until a beginner is fired or becomes qualified?

43. At a manufacturing plant, employees are classified as trainee (R), technician (T) or supervisor (S). Writing Q for an employee who quits we model their progress through the ranks as a Markov chain with transition probability

	R	T	S	Q
R	.2	.6	0	.2
T	0	.55	.15	.3
S	0	0	1	0
Q	0	0	0	1

(a) What fraction of recruits eventually make supervisor? (b) What is the expected time until a trainee quits or becomes supervisor?

44. The two previous problems have the following form:

	1	2	A	B
1	$1 - a - b$	a	0	b
2	0	$1 - c - d$	c	d
A	0	0	1	0
B	0	0	0	1

Show that (a) the probability of being absorbed in A rather than B is $ac/(a + b)(c + d)$ and (b) the expected time to absorption starting from 1 is $1/(a + b) + a/(a + b)(c + d)$.

45. The Markov chain associated with a manufacturing process may be described as follows: A part to be manufactured will begin the process by entering step 1. After step 1, 20% of the parts must be reworked, i.e., returned to step 1, 10% of the parts are thrown away, and 70% proceed to step 2. After step 2, 5% of the parts must be returned to the step 1, 10% to step 2, 5% are scrapped, and 80% emerge to be sold for a profit. (a) Formulate a four-state Markov chain with states 1, 2, 3, and 4 where 3 = a part that was scrapped and 4 = a part that was sold for a profit. (b) Compute the probability a part is scrapped in the production process.

46. Six children (Dick, Helen, Joni, Mark, Sam, and Tony) play catch. If Dick has the ball he is equally likely to throw it to Helen, Mark, Sam, and Tony. If Helen has the ball she is equally likely to throw it to Dick, Joni, Sam, and Tony. If Sam has the ball he is equally likely to throw it to Dick, Helen, Mark, and Tony. If either Joni or Tony gets the ball, they keep throwing it to each other. If Mark gets the ball he runs away with it. (a) Find the transition probability. (b) Suppose Dick has the ball at the beginning of the game. What is the probability Mark will end up with it?