Markov Chains

Sometimes we are interested in how a random variable changes over time. For example, we may want to know how the price of a share of stock or a firm's market share evolves. The study of how a random variable changes over time includes stochastic processes, which are explained in this chapter. In particular, we focus on a type of stochastic process known as a Markov chain. Markov chains have been applied in areas such as education, marketing, health services, finance, accounting, and production. We begin by defining the concept of a stochastic process. In the rest of the chapter, we will discuss the basic ideas needed for an understanding of Markov chains.

17.1 What Is a Stochastic Process?

Suppose we observe some characteristic of a system at discrete points in time (labeled 0, 1, 2, . . .). Let $\mathbf{X}_t$ be the value of the system characteristic at time t. In most situations, $\mathbf{X}_t$ is not known with certainty before time t and may be viewed as a random variable. A **discrete-time stochastic process** is simply a description of the relation between the random variables $\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2, \ldots$. Some examples of discrete-time stochastic processes follow.

EXAMPLE 1

At time 0, I have \$2. At times 1, 2, . . . , I play a game in which I bet \$1. With probability p, I win the game, and with probability $1 - p$, I lose the game. My goal is to increase my capital to \$4, and as soon as I do, the game is over. The game is also over if my capital is reduced to \$0. If we define $\mathbf{X}_t$ to be my capital position after the time t game (if any) is played, then $\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_t$ may be viewed as a discrete-time stochastic process. Note that $\mathbf{X}_0 = 2$ is a known constant, but $\mathbf{X}_1$ and later $\mathbf{X}_t$'s are random. For example, with probability p, $\mathbf{X}_1 = 3$, and with probability $1 - p$, $\mathbf{X}_1 = 1$. Note that if $\mathbf{X}_t = 4$, then $\mathbf{X}_{t+1}$ and all later $\mathbf{X}_t$'s will also equal 4. Similarly, if $\mathbf{X}_t = 0$, then $\mathbf{X}_{t+1}$ and all later $\mathbf{X}_t$'s will also equal 0. For obvious reasons, this type of situation is called a *gambler's ruin* problem.

EXAMPLE 2

An urn contains two unpainted balls at present. We choose a ball at random and flip a coin. If the chosen ball is unpainted and the coin comes up heads, we paint the chosen unpainted ball red; if the chosen ball is unpainted and the coin comes up tails, we paint the chosen unpainted ball black. If the ball has already been painted, then (whether heads or tails has been tossed) we change the color of the ball (from red to black or from black to red). To model this situation as a stochastic process, we define time t to be the time after the coin has been flipped for the tth time and the chosen ball has been painted. The state at any time may be described by the vector $[u \quad r \quad b]$, where u is the number of unpainted balls in the urn, r is the number of red balls in the urn, and b is the number of black balls in the urn. We are given that $\mathbf{X}_0 = [2 \quad 0 \quad 0]$. After the first coin toss, one ball will have been painted either red or black, and the state will be either $[1 \quad 1 \quad 0]$ or $[1 \quad 0 \quad 1]$. Hence, we can be sure that $\mathbf{X}_1 = [1 \quad 1 \quad 0]$ or $\mathbf{X}_1 = [1 \quad 0 \quad 1]$. Clearly, there must be some sort of relation between the $\mathbf{X}_t$'s. For example, if $\mathbf{X}_t = [0 \quad 2 \quad 0]$, we can be sure that $\mathbf{X}_{t+1}$ will be $[0 \quad 1 \quad 1]$.

EXAMPLE 3 **CSL Computer Stock**

Let $\mathbf{X}_0$ be the price of a share of CSL Computer stock at the beginning of the current trading day. Also, let $\mathbf{X}_t$ be the price of a share of CSL stock at the beginning of the tth trading day in the future. Clearly, knowing the values of $\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_t$ tells us something about the probability distribution of $\mathbf{X}_{t+1}$; the question is, what does the past (stock prices up to time t) tell us about $\mathbf{X}_{t+1}$? The answer to this question is of critical importance in finance. (See Section 17.2 for more details.)

We close this section with a brief discussion of continuous-time stochastic processes. A **continuous-time stochastic process** is simply a stochastic process in which the state of the system can be viewed at any time, not just at discrete instants in time. For example, the number of people in a supermarket t minutes after the store opens for business may be viewed as a continuous-time stochastic process. (Models involving continuous-time stochastic processes are studied in Chapter 20.) Since the price of a share of stock can be observed at any time (not just the beginning of each trading day), it may be viewed as a continuous-time stochastic process. Viewing the price of a share of stock as a continuous-time stochastic process has led to many important results in the theory of finance, including the famous Black–Scholes option pricing formula.

17.2 What Is a Markov Chain?

One special type of discrete-time stochastic process is called a *Markov chain.* To simplify our exposition, we assume that at any time, the discrete-time stochastic process can be in one of a finite number of states labeled $1, 2, \ldots, s$.

DEFINITION ■ A discrete-time stochastic process is a **Markov chain** if, for $t = 0, 1, 2, \ldots$ and all states,

$$P(\mathbf{X}_{t+1} = i_{t+1} | \mathbf{X}_t = i_t, \mathbf{X}_{t-1} = i_{t-1}, \ldots, \mathbf{X}_1 = i_1, \mathbf{X}_0 = i_0) = P(\mathbf{X}_{t+1} = i_{t+1} | \mathbf{X}_t = i_t) \quad ■ \tag{1}$$

Essentially, (1) says that the probability distribution of the state at time $t + 1$ depends on the state at time t (i_t) and does not depend on the states the chain passed through on the way to i_t at time t.

In our study of Markov chains, we make the further assumption that for all states i and j and all t, $P(\mathbf{X}_{t+1} = j | \mathbf{X}_t = i)$ is independent of t. This assumption allows us to write

$$P(\mathbf{X}_{t+1} = j | \mathbf{X}_t = i) = p_{ij} \tag{2}$$

where p_{ij} is the probability that given the system is in state i at time t, it will be in a state j at time $t + 1$. If the system moves from state i during one period to state j during the next period, we say that a **transition** from i to j has occurred. The p_{ij}'s are often referred to as the **transition probabilities** for the Markov chain.

Equation (2) implies that the probability law relating the next period's state to the current state does not change (or remains stationary) over time. For this reason, (2) is often called the **Stationarity Assumption.** Any Markov chain that satisfies (2) is called a **stationary Markov chain.**

Our study of Markov chains also requires us to define q_i to be the probability that the chain is in state i at time 0; in other words, $P(\mathbf{X}_0 = i) = q_i$. We call the vector $\mathbf{q} = [q_1 \quad q_2 \quad \cdots \quad q_s]$ the **initial probability distribution** for the Markov chain. In most applications, the transition probabilities are displayed as an $s \times s$ **transition probability matrix** P. The transition probability matrix P may be written as

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1s} \\ p_{21} & p_{22} & \cdots & p_{2s} \\ \vdots & \vdots & & \vdots \\ p_{s1} & p_{s2} & \cdots & p_{ss} \end{bmatrix}$$

Given that the state at time t is i, the process must be somewhere at time $t + 1$. This means that for each i,

$$\sum_{j=1}^{j=s} P(\mathbf{X}_{t+1} = j | P(\mathbf{X}_t = i)) = 1$$

$$\sum_{j=1}^{j=s} p_{ij} = 1$$

We also know that each entry in the P matrix must be nonnegative. Hence, all entries in the transition probability matrix are nonnegative, and the entries in each row must sum to 1.

EXAMPLE 1 The Gambler's Ruin (Continued)

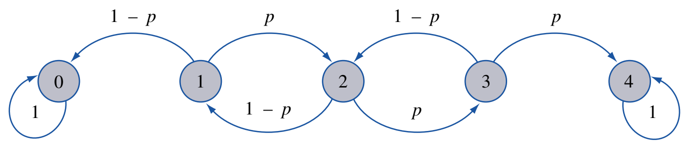
Find the transition matrix for Example 1.

Solution Since the amount of money I have after $t + 1$ plays of the game depends on the past history of the game only through the amount of money I have after t plays, we definitely have a Markov chain. Since the rules of the game don't change over time, we also have a stationary Markov chain. The transition matrix is as follows (state i means that we have i dollars):

$$P = \begin{matrix} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{matrix} \text{State} \\ \begin{matrix} \$0 & \$1 & \$2 & \$3 & \$4 \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1-p & 0 & p & 0 & 0 \\ 0 & 1-p & 0 & p & 0 \\ 0 & 0 & 1-p & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix}$$

If the state is \$0 or \$4, I don't play the game anymore, so the state cannot change; hence, $p_{00} = p_{44} = 1$. For all other states, we know that with probability p, the next period's state will exceed the current state by 1, and with probability $1 - p$, the next period's state will be 1 less than the current state.

FIGURE 1
Graphical Representation of Transition Matrix for Gambler's Ruin



A transition matrix may be represented by a graph in which each node represents a state and arc (i, j) represents the transition probability p_{ij}. Figure 1 gives a graphical representation of Example 1's transition probability matrix.

EXAMPLE 2 Choosing Balls (Continued)

Find the transition matrix for Example 2.

Solution Since the state of the urn after the next coin toss only depends on the past history of the process through the state of the urn after the current coin toss, we have a Markov chain. Since the rules don't change over time, we have a stationary Markov chain. The transition matrix for Example 2 is as follows:

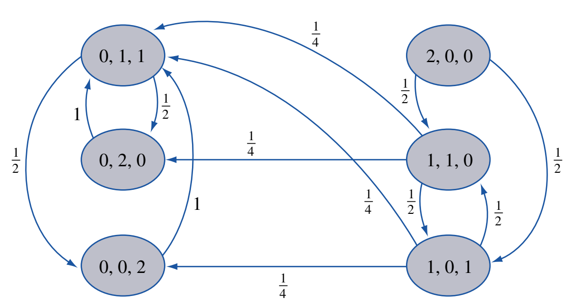
$$P = \begin{array}{c} \\ [0\ \ 1\ \ 1] \\ [0\ \ 2\ \ 0] \\ [0\ \ 0\ \ 2] \\ [2\ \ 0\ \ 0] \\ [1\ \ 1\ \ 0] \\ [1\ \ 0\ \ 1] \end{array} \begin{array}{c} \text{State} \\ \begin{array}{cccccc} [0\ \ 1\ \ 1] & [0\ \ 2\ \ 0] & [0\ \ 0\ \ 2] & [2\ \ 0\ \ 0] & [1\ \ 1\ \ 0] & [1\ \ 0\ \ 1] \end{array} \\ \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{2} & 0 \end{bmatrix} \end{array}$$

To illustrate the determination of the transition matrix, we determine the [1 1 0] row of this transition matrix. If the current state is [1 1 0], then one of the events shown in Table 1 must occur. Thus, the next state will be [1 0 1] with probability $\frac{1}{2}$, [0 2 0] with probability $\frac{1}{4}$, and [0 1 1] with probability $\frac{1}{4}$. Figure 2 gives a graphical representation of this transition matrix.

TABLE 1
Computations of Transition Probabilities If Current State Is [1 1 0]

Event	Probability	New State
Flip heads and choose unpainted ball	$\frac{1}{4}$	[0 2 0]
Choose red ball	$\frac{1}{2}$	[1 0 1]
Flip tails and choose unpainted ball	$\frac{1}{4}$	[0 1 1]

FIGURE 2
Graphical Representation of Transition Matrix for Urn



EXAMPLE 3 **CSL Computer Stock (Continued)**

In recent years, students of finance have devoted much effort to answering the question of whether the daily price of a stock share can be described by a Markov chain. Suppose the daily price of a stock share (such as CSL Computer stock) can be described by a Markov chain. What does that tell us? Simply that the probability distribution of tomorrow's price for one share of CSL stock depends *only* on today's price of CSL stock, *not* on the past prices of CSL stock. If the price of a stock share can be described by a Markov chain, the "chartists" who attempt to predict future stock prices on the basis of the patterns followed by past stock prices are barking up the wrong tree. For example, suppose the daily price of a share of CSL stock follows a Markov chain, and today's price for a share of CSL stock is $50. Then to predict tomorrow's price of a share of CSL stock, it does not matter whether the price has increased or decreased during each of the last 30 days. In either situation (or any other situation that might have led to today's $50 price), a prediction of tomorrow's stock price should be based only on the fact that today's price of CSL stock is $50. At this time, the consensus is that for most stocks the daily price of the stock can be described as a Markov chain. This idea is often referred to as the **efficient market hypothesis.**

PROBLEMS

Group A

1 In Smalltown, 90% of all sunny days are followed by sunny days, and 80% of all cloudy days are followed by cloudy days. Use this information to model Smalltown's weather as a Markov chain.

2 Consider an inventory system in which the sequence of events during each period is as follows. (1) We observe the inventory level (call it i) at the beginning of the period. (2) If $i \leq 1$, $4 - i$ units are ordered. If $i \geq 2$, 0 units are ordered. Delivery of all ordered units is immediate. (3) With probability $\frac{1}{3}$, 0 units are demanded during the period; with probability $\frac{1}{3}$, 1 unit is demanded during the period; and with probability $\frac{1}{3}$, 2 units are demanded during the period. (4) We observe the inventory level at the beginning of the next period.

Define a period's state to be the period's beginning inventory level. Determine the transition matrix that could be used to model this inventory system as a Markov chain.

3 A company has two machines. During any day, each machine that is working at the beginning of the day has a $\frac{1}{3}$ chance of breaking down. If a machine breaks down during the day, it is sent to a repair facility and will be working two days after it breaks down. (Thus, if a machine breaks down during day 3, it will be working at the beginning of day 5.) Letting the state of the system be the number of machines working at the beginning of the day, formulate a transition probability matrix for this situation.

Group B

4 Referring to Problem 1, suppose that tomorrow's Smalltown weather depends on the last two days of Smalltown weather, as follows: (1) If the last two days have been sunny, then 95% of the time, tomorrow will be sunny. (2) If yesterday was cloudy and today is sunny, then 70% of the time, tomorrow will be sunny. (3) If yesterday was sunny and today is cloudy, then 60% of the time, tomorrow will be cloudy. (4) If the last two days have been cloudy, then 80% of the time, tomorrow will be cloudy.

Using this information, model Smalltown's weather as a Markov chain. If tomorrow's weather depended on the last three days of Smalltown weather, how many states will be needed to model Smalltown's weather as a Markov chain? (*Note:* The approach used in this problem can be used to model a discrete-time stochastic process as a Markov chain even if $\mathbf{X}_{t+1}$ depends on states prior to $\mathbf{X}_t$, such as $\mathbf{X}_{t-1}$ in the current example.)

5 Let $\mathbf{X}_t$ be the location of your token on the Monopoly board after t dice rolls. Can $\mathbf{X}_t$ be modeled as a Markov chain? If not, how can we modify the definition of the state at time t so that $\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_t, \ldots$ would be a Markov chain? (*Hint:* How does a player go to Jail? In this problem, assume that players who are sent to Jail stay there until they roll doubles or until they have spent three turns there, whichever comes first.)

6 In Problem 3, suppose a machine that breaks down returns to service three days later (for instance, a machine that breaks down during day 3 would be back in working order at the beginning of day 6). Determine a transition probability matrix for this situation.

17.3 *n*-Step Transition Probabilities

Suppose we are studying a Markov chain with a known transition probability matrix P. (Since all chains that we will deal with are stationary, we will not bother to label our Markov chains as stationary.) A question of interest is: If a Markov chain is in state i at time m, what is the probability that n periods later the Markov chain will be in state j? Since we are dealing with a stationary Markov chain, this probability will be independent of m, so we may write

$$P(\mathbf{X}_{m+n} = j | \mathbf{X}_m = i) = P(\mathbf{X}_n = j | \mathbf{X}_0 = i) = P_{ij}(n)$$

where $P_{ij}(n)$ is called the ***n*-step probability** of a transition from state i to state j.

Clearly, $P_{ij}(1) = p_{ij}$. To determine $P_{ij}(2)$, note that if the system is now in state i, then for the system to end up in state j two periods from now, we must go from state i to some state k and then go from state k to state j (see Figure 3). This reasoning allows us to write

$$P_{ij}(2) = \sum_{k=1}^{k=s} (\text{probability of transition from } i \text{ to } k) \times (\text{probability of transition from } k \text{ to } j)$$

Using the definition of P, the transition probability matrix, we rewrite the last equation as

$$P_{ij}(2) = \sum_{k=1}^{k=s} p_{ik} p_{kj} \tag{3}$$

The right-hand side of (3) is just the scalar product of row i of the P matrix with column j of the P matrix. Hence, $P_{ij}(2)$ is the ijth element of the matrix P^2. By extending this reasoning, it can be shown that for $n > 1$,

$$P_{ij}(n) = ij\text{th element of } P^n \tag{4}$$

Of course, for $n = 0$, $P_{ij}(0) = P(\mathbf{X}_0 = j | \mathbf{X}_0 = i)$, so we must write

$$P_{ij}(0) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}$$

We illustrate the use of Equation (4) in Example 4.

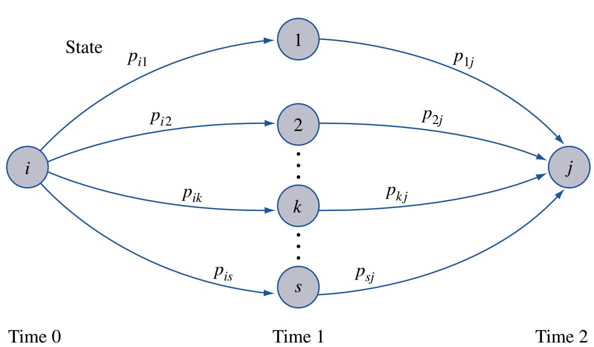


FIGURE 3
$P_{ij}(2) = p_{i1}p_{1j} + p_{i2}p_{2j} + \cdots + p_{is}p_{sj}$

EXAMPLE 4 The Cola Example

Suppose the entire cola industry produces only two colas. Given that a person last purchased cola 1, there is a 90% chance that her next purchase will be cola 1. Given that a person last purchased cola 2, there is an 80% chance that her next purchase will be cola 2.

1 If a person is currently a cola 2 purchaser, what is the probability that she will purchase cola 1 two purchases from now?

2 If a person is currently a cola 1 purchaser, what is the probability that she will purchase cola 1 three purchases from now?

Solution We view each person's purchases as a Markov chain with the state at any given time being the type of cola the person last purchased. Hence, each person's cola purchases may be represented by a two-state Markov chain, where

State 1 = person has last purchased cola 1

State 2 = person has last purchased cola 2

If we define $\mathbf{X}_n$ to be the type of cola purchased by a person on her nth future cola purchase (present cola purchase = $\mathbf{X}_0$), then $\mathbf{X}_0, \mathbf{X}_1, \ldots$ may be described as the Markov chain with the following transition matrix:

$$P = \begin{array}{c} \\ \text{Cola 1} \\ \text{Cola 2} \end{array} \begin{array}{c} \begin{array}{cc} \text{Cola 1} & \text{Cola 2} \end{array} \\ \begin{bmatrix} .90 & .10 \\ .20 & .80 \end{bmatrix} \end{array}$$

We can now answer questions 1 and 2.

1 We seek $P(\mathbf{X}_2 = 1|\mathbf{X}_0 = 2) = P_{21}(2) = $ element 21 of P^2:

$$P^2 = \begin{bmatrix} .90 & .10 \\ .20 & .80 \end{bmatrix} \begin{bmatrix} .90 & .10 \\ .20 & .80 \end{bmatrix} = \begin{bmatrix} .83 & .17 \\ .34 & .66 \end{bmatrix}$$

Hence, $P_{21}(2) = .34$. This means that the probability is .34 that two purchases in the future a cola 2 drinker will purchase cola 1. By using basic probability theory, we may obtain this answer in a different way (see Figure 4). Note that $P_{21}(2) =$ (probability that next purchase is cola 1 and second purchase is cola 1) + (probability that next purchase is cola 2 and second purchase is cola 1) $= p_{21}p_{11} + p_{22}p_{21} = (.20)(.90) + (.80)(.20) = .34$.

2 We seek $P_{11}(3) = $ element 11 of P^3:

$$P^3 = P(P^2) = \begin{bmatrix} .90 & .10 \\ .20 & .80 \end{bmatrix} \begin{bmatrix} .83 & .17 \\ .34 & .66 \end{bmatrix} = \begin{bmatrix} .781 & .219 \\ .438 & .562 \end{bmatrix}$$

Therefore, $P_{11}(3) = .781$.

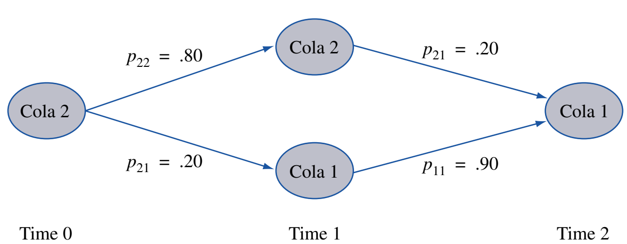


FIGURE 4 Probability That Two Periods from Now, a Cola 2 Purchaser Will Purchase Cola 1 Is .20(.90) + .80(.20) = .34

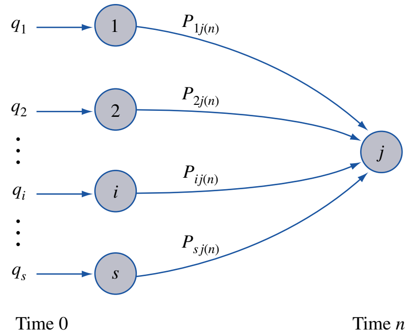


FIGURE 5
Determination of Probability of Being in State j at Time n When Initial State Is Unknown

In many situations, we do not know the state of the Markov chain at time 0. As defined in Section 17.2, let q_i be the probability that the chain is in state i at time 0. Then we can determine the probability that the system is in state i at time n by using the following reasoning (see Figure 5).

$$\begin{aligned}
&\text{Probability of being in state } j \text{ at time } n \\
&= \sum_{i=1}^{i=s} (\text{probability that state is originally } i) \\
&\quad \times (\text{probability of going from } i \text{ to } j \text{ in } n \text{ transitions}) \\
&= \sum_{i=1}^{i=s} q_i P_{ij}(n) \\
&= \mathbf{q}(\text{column } j \text{ of } P^n)
\end{aligned} \qquad (5)$$

where $\mathbf{q} = [q_1 \quad q_2 \quad \cdots \quad q_s]$.

To illustrate the use of (5), we answer the following question: Suppose 60% of all people now drink cola 1, and 40% now drink cola 2. Three purchases from now, what fraction of all purchasers will be drinking cola 1? Since $\mathbf{q} = [.60 \quad .40]$ and $\mathbf{q}$(column 1 of P^3) = probability that three purchases from now a person drinks cola 1, the desired probability is

$$[.60 \quad .40]\begin{bmatrix} .781 \\ .438 \end{bmatrix} = .6438$$

Hence, three purchases from now, 64% of all purchasers will be purchasing cola 1.

To illustrate the behavior of the n-step transition probabilities for large values of n, we have computed several of the n-step transition probabilities for the Cola example in Table 2.

TABLE 2
***n*-Step Transition Probabilities for Cola Drinkers**

n	$P_{11}(n)$	$P_{12}(n)$	$P_{21}(n)$	$P_{22}(n)$
1	.90	.10	.20	.80
2	.83	.17	.34	.66
3	.78	.22	.44	.56
4	.75	.25	.51	.49
5	.72	.28	.56	.44
10	.68	.32	.65	.35
20	.67	.33	.67	.33
30	.67	.33	.67	.33
40	.67	.33	.67	.33

For large n, both $P_{11}(n)$ and $P_{21}(n)$ are nearly constant and approach .67. This means that for large n, no matter what the initial state, there is a .67 chance that a person will be a cola 1 purchaser. Similarly, we see that for large n, both $P_{12}(n)$ and $P_{22}(n)$ are nearly constant and approach .33. This means that for large n, no matter what the initial state, there is a .33 chance that a person will be a cola 2 purchaser. In Section 5.5, we make a thorough study of this settling down of the n-step transition probabilities.

REMARK We can easily multiply matrices on a spreadsheet using the MMULT command, as discussed in Section 13.7.

PROBLEMS

Group A

1 Each American family is classified as living in an urban, rural, or suburban location. During a given year, 15% of all urban families move to a suburban location, and 5% move to a rural location; also, 6% of all suburban families move to an urban location, and 4% move to a rural location; finally, 4% of all rural families move to an urban location, and 6% move to a suburban location.

a If a family now lives in an urban location, what is the probability that it will live in an urban area two years from now? A suburban area? A rural area?

b Suppose that at present, 40% of all families live in an urban area, 35% live in a suburban area, and 25% live in a rural area. Two years from now, what percentage of American families will live in an urban area?

c What problems might occur if this model were used to predict the future population distribution of the United States?

2 The following questions refer to Example 1.

a After playing the game twice, what is the probability that I will have \$3? How about \$2?

b After playing the game three times, what is the probability that I will have \$2?

3 In Example 2, determine the following n-step transition probabilities:

a After two balls are painted, what is the probability that the state is [0 2 0]?

b After three balls are painted, what is the probability that the state is [0 1 1]? (Draw a diagram like Figure 4.)

17.4 Classification of States in a Markov Chain

In Section 17.3, we mentioned the fact that after many transitions, the n-step transition probabilities tend to settle down. Before we can discuss this in more detail, we need to study how mathematicians classify the states of a Markov chain. We use the following transition matrix to illustrate most of the following definitions (see Figure 6).

$$P = \begin{bmatrix} .4 & .6 & 0 & 0 & 0 \\ .5 & .5 & 0 & 0 & 0 \\ 0 & 0 & .3 & .7 & 0 \\ 0 & 0 & .5 & .4 & .1 \\ 0 & 0 & 0 & .8 & .2 \end{bmatrix}$$

DEFINITION ■ Given two states i and j, a **path** from i to j is a sequence of transitions that begins in i and ends in j, such that each transition in the sequence has a positive probability of occurring. ■

A state j is **reachable** from state i if there is a path leading from i to j. ■

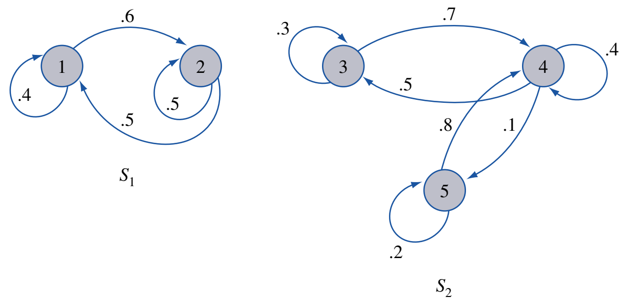


FIGURE 6 Graphical Representation of Transition Matrix

DEFINITION ■ Two states i and j are said to **communicate** if j is reachable from i, and i is reachable from j. ■

For the transition probability matrix P represented in Figure 6, state 5 is reachable from state 3 (via the path 3–4–5), but state 5 is not reachable from state 1 (there is no path from 1 to 5 in Figure 6). Also, states 1 and 2 communicate (we can go from 1 to 2 and from 2 to 1).

DEFINITION ■ A set of states S in a Markov chain is a **closed set** if no state outside of S is reachable from any state in S. ■

From the Markov chain with transition matrix P in Figure 6, $S_1 = \{1, 2\}$ and $S_2 = \{3, 4, 5\}$ are both closed sets. Observe that once we enter a closed set, we can never leave the closed set (in Figure 6, no arc begins in S_1 and ends in S_2 or begins in S_2 and ends in S_1).

DEFINITION ■ A state i is an **absorbing state** if $p_{ii} = 1$. ■

Whenever we enter an absorbing state, we never leave the state. In Example 1, the gambler's ruin, states 0 and 4 are absorbing states. Of course, an absorbing state is a closed set containing only one state.

DEFINITION ■ A state i is a **transient state** if there exists a state j that is reachable from i, but the state i is not reachable from state j. ■

In other words, a state i is transient if there is a way to leave state i that never returns to state i. In the gambler's ruin example, states 1, 2, and 3 are transient states. For example (see Figure 1), from state 2, it is possible to go along the path 2–3–4, but there is no way to return to state 2 from state 4. Similarly, in Example 2, [2 0 0], [1 1 0], and [1 0 1] are all transient states (in Figure 2, there is a path from [1 0 1] to [0 0 2], but once both balls are painted, there is no way to return to [1 0 1]).

After a large number of periods, the probability of being in any transient state i is zero. Each time we enter a transient state i, there is a positive probability that we will leave i forever and end up in the state j described in the definition of a transient state. Thus, eventually we are sure to enter state j (and then we will never return to state i). To illustrate, in Example 2, suppose we are in the transient state [1 0 1]. With probability 1, the unpainted ball will eventually be painted, and we will never reenter state [1 0 1] (see Figure 2).

DEFINITION ■ If a state is not transient, it is called a **recurrent state.** ■

In Example 1, states 0 and 4 are recurrent states (and also absorbing states), and in Example 2, [0 2 0], [0 0 2], and [0 1 1] are recurrent states. For the transition matrix P in Figure 6, all states are recurrent.

DEFINITION ■ A state i is **periodic** with period $k > 1$ if k is the smallest number such that all paths leading from state i back to state i have a length that is a multiple of k. If a recurrent state is not periodic, it is referred to as **aperiodic.** ■

For the Markov chain with transition matrix

$$Q = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

each state has period 3. For example, if we begin in state 1, the only way to return to state 1 is to follow the path 1–2–3–1 for some number of times (say, m). (See Figure 7.) Hence, any return to state 1 will take $3m$ transitions, so state 1 has period 3. Wherever we are, we are sure to return three periods later.

DEFINITION ■ If all states in a chain are recurrent, aperiodic, and communicate with each other, the chain is said to be **ergodic.** ■

The gambler's ruin example is not an ergodic chain, because (for example) states 3 and 4 do not communicate. Example 2 is also not an ergodic chain, because (for example) [2 0 0] and [0 1 1] do not communicate. Example 4, the cola example, is an ergodic Markov chain. Of the following three Markov chains, P_1 and P_3 are ergodic, and P_2 is not ergodic.

$$P_1 = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{3}{4} \end{bmatrix} \qquad \text{Ergodic}$$

$$P_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{4} & \frac{3}{4} \end{bmatrix} \qquad \text{Nonergodic}$$

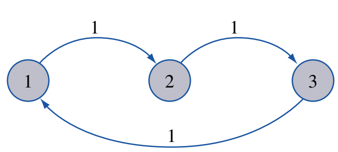


FIGURE 7
A Periodic Markov Chain $k = 3$

$$P_3 = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \end{bmatrix} \qquad \text{Ergodic}$$

P_2 is not ergodic because there are two closed classes of states (class 1 = {1, 2} and class 2 = {3, 4}), and the states in different classes do not communicate with each other.

After the next two sections, the importance of the concepts introduced in this section will become clear.

PROBLEMS

Group A

1 In Example 1, what is the period of states 1 and 3?

2 Is the Markov chain of Section 17.3, Problem 1, an ergodic Markov chain?

3 Consider the following transition matrix:

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{2}{3} \end{bmatrix}$$

a Which states are transient?

b Which states are recurrent?

c Identify all closed sets of states.

d Is this chain ergodic?

4 For each of the following chains, determine whether the Markov chain is ergodic. Also, for each chain, determine the recurrent, transient, and absorbing states.

$$P_1 = \begin{bmatrix} 0 & .8 & .2 \\ .3 & .7 & 0 \\ .4 & .5 & .1 \end{bmatrix} \qquad P_2 = \begin{bmatrix} .2 & .8 & 0 & 0 \\ 0 & 0 & .9 & .1 \\ .4 & .5 & .1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

5 Fifty-four players (including Gabe Kaplan and James Garner) participated in the 1980 World Series of Poker. Each player began with $10,000. Play continued until one player had won everybody else's money. If the World Series of Poker were to be modeled as a Markov chain, how many absorbing states would the chain have?

6 Which of the following chains is ergodic?

$$P_1 = \begin{bmatrix} .4 & 0 & .6 \\ .3 & .3 & .4 \\ 0 & .5 & .5 \end{bmatrix} \qquad P_2 = \begin{bmatrix} .7 & 0 & 0 & .3 \\ .2 & .2 & .4 & .2 \\ .6 & .1 & .1 & .2 \\ .2 & 0 & 0 & .8 \end{bmatrix}$$

17.5 Steady-State Probabilities and Mean First Passage Times

In our discussion of the cola example (Example 4), we found that after a long time, the probability that a person's next cola purchase would be cola 1 approached .67 and .33 that it would be cola 2 (see Table 2). These probabilities *did not* depend on whether the person was initially a cola 1 or a cola 2 drinker. In this section, we discuss the important concept of steady-state probabilities, which can be used to describe the long-run behavior of a Markov chain.

The following result is vital to an understanding of steady-state probabilities and the long-run behavior of Markov chains.

THEOREM 1

Let P be the transition matrix for an s-state ergodic chain.† Then there exists a vector $\pi = [\pi_1 \quad \pi_2 \quad \cdots \quad \pi_s]$ such that

†To see why Theorem 1 fails to hold for a nonergodic chain, see Problems 11 and 12 at the end of this section. For a proof of this theorem, see Isaacson and Madsen (1976, Chapter 3).

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

Recall that the ijth element of P^n is $P_{ij}(n)$. Theorem 1 tells us that for any initial state i,

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

Observe that for large n, P^n approaches a matrix with identical rows. This means that after a long time, the Markov chain settles down, and (independent of the initial state i) there is a probability π_j that we are in state j.

The vector $\pi = [\pi_1 \quad \pi_2 \quad \cdots \quad \pi_s]$ is often called the **steady-state distribution,** or **equilibrium distribution,** for the Markov chain. For a given chain with transition matrix P, how can we find the steady-state probability distribution? From Theorem 1, observe that for large n and all i,

$$P_{ij}(n+1) \cong P_{ij}(n) \cong \pi_j \tag{6}$$

Since $P_{ij}(n+1) = (\text{row } i \text{ of } P^n)(\text{column } j \text{ of } P)$, we may write

$$P_{ij}(n+1) = \sum_{k=1}^{k=s} P_{ik}(n)p_{kj} \tag{7}$$

If n is large, substituting (6) into (7) yields

$$\pi_j = \sum_{k=1}^{k=s} \pi_k p_{kj} \tag{8}$$

In matrix form, (8) may be written as

$$\pi = \pi P \tag{8'}$$

Unfortunately, the system of equations specified in (8) has an infinite number of solutions, because the rank of the P matrix always turns out to be $\leq s - 1$ (see Chapter 2, Review Problem 21). To obtain unique values of the steady-state probabilities, note that for any n and any i,

$$P_{i1}(n) + P_{i2}(n) + \cdots + P_{is}(n) = 1 \tag{9}$$

Letting n approach infinity in (9), we obtain

$$\pi_1 + \pi_2 + \cdots + \pi_s = 1 \tag{10}$$

Thus, after replacing any of the equations in (8) with (10), we may use (8) to solve for the steady-state probabilities.

To illustrate how to find the steady-state probabilities, we find the steady-state probabilities for Example 4, the cola example. Recall that the transition matrix for Example 4 was

$$P = \begin{bmatrix} .90 & .10 \\ .20 & .80 \end{bmatrix}$$

Then (8) or (8′) yields

$$[\pi_1 \quad \pi_2] = [\pi_1 \quad \pi_2]\begin{bmatrix} .90 & .10 \\ .20 & .80 \end{bmatrix}$$

$$\pi_1 = .90\pi_1 + .20\pi_2$$

$$\pi_2 = .10\pi_1 + .80\pi_2$$

Replacing the second equation with the condition $\pi_1 + \pi_2 = 1$, we obtain the system

$$\pi_1 = .90\pi_1 + .20\pi_2$$

$$1 = \pi_1 + \pi_2$$

Solving for π_1 and π_2 we obtain $\pi_1 = \frac{2}{3}$ and $\pi_2 = \frac{1}{3}$. Hence, after a long time, there is a $\frac{2}{3}$ probability that a given person will purchase cola 1 and a $\frac{1}{3}$ probability that a given person will purchase cola 2.

Transient Analysis

A glance at Table 2 shows that for Example 4, the steady state is reached (to two decimal places) after only ten transitions. No general rule can be given about how quickly a Markov chain reaches the steady state, but if P contains very few entries that are near 0 or near 1, the steady state is usually reached very quickly. The behavior of a Markov chain before the steady state is reached is often called **transient** (or short-run) **behavior.** To study the transient behavior of a Markov chain, one simply uses the formulas for $P_{ij}(n)$ given in (4) and (5). It's nice to know, however, that for large n, the steady-state probabilities accurately describe the probability of being in any state.

Intuitive Interpretation of Steady-State Probabilities

An intuitive interpretation can be given to the steady-state probability equations (8). By subtracting $\pi_j p_{jj}$ from both sides of (8), we obtain

$$\pi_j(1 - p_{jj}) = \sum_{k \neq j} \pi_k p_{kj} \tag{11}$$

Equation (11) states that in the steady state,

$$\begin{aligned} &\text{Probability that a particular transition leaves state } j \\ &\quad = \text{probability that a particular transition enters state } j \end{aligned} \tag{12}$$

Recall that in the steady state, the probability that the system is in state j is π_j. From this observation, it follows that

$$\begin{aligned} &\text{Probability that a particular transition leaves state } j \\ &\quad = (\text{probability that the current period begins in } j) \\ &\quad \times (\text{probability that the current transition leaves } j) \\ &\quad = \pi_j(1 - p_{jj}) \end{aligned}$$

and

$$\begin{aligned} &\text{Probability that a particular transition enters state } j \\ &\quad = \sum_k (\text{probability that the current period begins in } k \neq j) \\ &\quad \times (\text{probability that the current transition enters } j) \end{aligned}$$

$$= \sum_{k \neq j} \pi_k p_{kj}$$

Equation (11) is reasonable; if (11) were violated for any state, then for some state j, the right-hand side of (11) would exceed the left-hand side of (11). This would result in probability "piling up" at state j, and a steady-state distribution would not exist. Equation (11) may be viewed as saying that in the steady state, the "flow" of probability into each state must equal the flow of probability out of each state. This explains why steady-state probabilities are often called equilibrium probabilities.

Use of Steady-State Probabilities in Decision Making

EXAMPLE 5 The Cola Example (Continued)

In Example 4, suppose that each customer makes one purchase of cola during any week (52 weeks = 1 year). Suppose there are 100 million cola customers. One selling unit of cola costs the company \$1 to produce and is sold for \$2. For \$500 million per year, an advertising firm guarantees to decrease from 10% to 5% the fraction of cola 1 customers who switch to cola 2 after a purchase. Should the company that makes cola 1 hire the advertising firm?

Solution At present, a fraction $\pi_1 = \frac{2}{3}$ of all purchases are cola 1 purchases. Each purchase of cola 1 earns the company a \$1 profit. Since there are a total of 52(100,000,000), or 5.2 billion, cola purchases each year, the cola 1 company's current annual profit is

$$\tfrac{2}{3}(5{,}200{,}000{,}000) = \$3{,}466{,}666{,}667$$

The advertising firm is offering to change the P matrix to

$$P_1 = \begin{bmatrix} .95 & .05 \\ .20 & .80 \end{bmatrix}$$

For P_1, the steady-state equations become

$$\pi_1 = .95\pi_1 + .20\pi_2$$
$$\pi_2 = .05\pi_1 + .80\pi_2$$

Replacing the second equation by $\pi_1 + \pi_2 = 1$ and solving, we obtain $\pi_1 = .8$ and $\pi_2 = .2$. Now the cola 1 company's annual profit will be

$$(.80)(5{,}200{,}000{,}000) - 500{,}000{,}000 = \$3{,}660{,}000{,}000$$

Hence, the cola 1 company should hire the ad agency.

EXAMPLE 6 Playing Monopoly

With the assumption that each Monopoly player who goes to Jail stays until he or she rolls doubles or has spent three turns in Jail, the steady-state probability of a player landing on any Monopoly square has been determined by Ash and Bishop (1972) (see Table 3).[†] These steady-state probabilities can be used to measure the cost-effectiveness of various monopolies. For example, it costs \$1,500 to build hotels on the Orange monopoly. Each time a player lands on a Tennessee Ave. or a St. James Place hotel, the owner of the monopoly receives \$950, and each time a player lands on a New York Ave. hotel, the owner

[†]This example is based on Ash and Bishop (1972).

TABLE 3
Steady-State Probabilities for Monopoly

	n Position	Steady-State Probability
0	Go	.0346
1	Mediterranean Ave.	.0237
2	Community Chest 1	.0218
3	Baltic Ave.	.0241
4	Income tax	.0261
5	Reading RR	.0332
6	Oriental Ave.	.0253
7	Chance 1	.0096
8	Vermont Ave.	.0258
9	Connecticut Ave.	.0237
10	Visiting jail	.0254
11	St. Charles Place	.0304
12	Electric Co.	.0311
13	State Ave.	.0258
14	Virginia Ave.	.0288
15	Pennsylvania RR	.0313
16	St. James Place	.0318
17	Community Chest 2	.0272
18	Tennessee Ave.	.0335
19	New York Ave.	.0334
20	Free parking	.0336
21	Kentucky Ave.	.0310
22	Chance 2	.0125
23	Indiana Ave.	.0305
24	Illinois Ave.	.0355
25	B and O RR	.0344
26	Atlantic Ave.	.0301
27	Ventnor Ave.	.0299
28	Water works	.0315
29	Marvin Gardens	.0289
30	Jail	.1123
31	Pacific Ave.	.0300
32	North Carolina Ave.	.0294
33	Community Chest 3	.0263
34	Pennsylvania Ave.	.0279
35	Short Line RR	.0272
36	Chance 3	.0096
37	Park Place	.0245
38	Luxury tax	.0295
39	Boardwalk	.0295

Source: Reprinted by permission from R. Ash and R. Bishop, "Monopoly as a Markov Process," *Mathematics Magazine* 45(1972):26–29. Copyright © 1972 Mathematical Association of America.

receives $1,000. From Table 3, we can compute the expected rent per dice roll earned by the Orange monopoly:

$$950(.0335) + 950(.0318) + 1{,}000(.0334) = \$95.44$$

Thus, per dollar invested, the Orange monopoly yields $\frac{95.44}{1,500} = \$0.064$ per dice roll.

Now let's consider the Green monopoly. To put hotels on the Green monopoly costs \$3,000. If a player lands on a North Carolina Ave. or a Pacific Ave. hotel, the owner receives \$1,275. If a player lands on a Pennsylvania Ave. hotel, the owner receives \$1,400. From Table 3, the average revenue per dice roll earned from hotels on the Green monopoly is

$$1{,}275(.0294) + 1{,}275(.0300) + 1{,}400(.0279) = \$114.80$$

Thus, per dollar invested, the Green monopoly yields only $\frac{114.80}{3,000} = \$0.038$ per dice roll.

This analysis shows that the Orange monopoly is superior to the Green monopoly. By the way, why does the Orange get landed on so often?

Mean First Passage Times

For an ergodic chain, let m_{ij} = expected number of transitions before we first reach state j, given that we are currently in state i; m_{ij} is called the **mean first passage time** from state i to state j. In Example 4, m_{12} would be the expected number of bottles of cola purchased by a person who just bought cola 1 before first buying a bottle of cola 2. Assume that we are currently in state i. Then with probability p_{ij}, it will take one transition to go from state i to state j. For $k \neq j$, we next go with probability p_{ik} to state k. In this case, it will take an average of $1 + m_{kj}$ transitions to go from i to j. This reasoning implies that

$$m_{ij} = p_{ij}(1) + \sum_{k \neq j} p_{ik}(1 + m_{kj})$$

Since

$$p_{ij} + \sum_{k \neq j} p_{ik} = 1$$

we may rewrite the last equation as

$$m_{ij} = 1 + \sum_{k \neq j} p_{ik} m_{kj} \tag{13}$$

By solving the linear equations given in (13), we may find all the mean first passage times. It can be shown that

$$m_{ii} = \frac{1}{\pi_i}$$

This can simplify the use of (13).

To illustrate the use of (13), let's solve for the mean first passage times in Example 4. Recall that $\pi_1 = \frac{2}{3}$ and $\pi_2 = \frac{1}{3}$. Then

$$m_{11} = \frac{1}{\ } = 1.5 \qquad \text{and} \qquad m_{22} = \frac{1}{\ } = 3$$

Markov.lng

Now (13) yields the following two equations:

$$m_{12} = 1 + p_{11}m_{12} = 1 + 0.9m_{12}, \qquad m_{21} = 1 + p_{22}m_{21} = 1 + 0.8m_{21}$$

Solving these two equations, we find that $m_{12} = 10$ and $m_{21} = 5$. This means, for example, that a person who last drank cola 1 will drink an average of ten bottles of soda before switching to cola 2.

Solving for Steady-State Probabilities and Mean First Passage Times on the Computer

Since we solve for steady-state probabilities and mean first passage times by solving a system of linear equations, we may use LINDO to determine them. Simply type in an objective function of 0, and type the equations you need to solve as your constraints.

Alternatively, you may use the following LINGO model (file Markov.lng) to determine steady-state probabilities and mean first passage times for an ergodic chain.

```
MODEL:
 1]
 2]SETS:
 3]STATE/1..2/:PI;
 4]SXS(STATE,STATE):TPROB,MFP;
 5]ENDSETS
 6]DATA:
 7]TPROB = .9,.1,
 8].2,.8;
 9]ENDDATA
 10]@FOR(STATE(J)|J #LT# @SIZE(STATE):
 11]PI(J) = @SUM(SXS(I,J): PI(I) * TPROB(I,J)););
 12]@SUM(STATE:PI) = 1;
 13]@FOR(SXS(I,J):MFP(I,J)=
 14]1+@SUM(STATE(K)|K#NE#J:TPROB(I,K)*MFP(K,J)););
END
```

In line 3, we define the set of states and associate a steady-state probability ($PI(I)$) with each state I. In line 4, we create for each pairing of states (I, J) a transition probability (TPROB(I, J)) which equals p_{ij} and MFP(I, J) which equals m_{ij}. The transition probabilities for the cola example are input in lines 7 and 8. In lines 10 and 11, we create (for each state except the highest-numbered state) the steady-state equation

$$PI(J) = \sum_{I} PI(I) * \text{TPROB}(I, J)$$

In line 12, we ensure that the steady-state probabilities sum to 1. In lines 13 and 14, we create the equations that must be solved to compute the mean first passage times. For each (I, J), lines 13–14 create the equation

$$\text{MFP}(I, J) = 1 + \sum_{K \neq J} \text{TPROB}(I, K) * \text{MFP}(K, J)$$

which is needed to compute the mean first passage times.

PROBLEMS

Group A

1 Find the steady-state probabilities for Problem 1 of Section 17.3.

2 For the gambler's ruin problem (Example 1), why is it unreasonable to talk about steady-state probabilities?

3 For each of the following Markov chains, determine the long-run fraction of the time that each state will be occupied.

a $\begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ **b** $\begin{bmatrix} .8 & .2 & 0 \\ 0 & .2 & .8 \\ .8 & .2 & 0 \end{bmatrix}$

c Find all mean first passage times for part (b).

4 At the beginning of each year, my car is in good, fair, or broken-down condition. A good car will be good at the beginning of next year with probability .85; fair with probability .10; or broken-down with probability .05. A fair car will be fair at the beginning of the next year with probability .70 or broken-down with probability .30. It costs $6,000 to purchase a good car; a fair car can be traded in for $2,000; and a broken-down car has no trade-in value and must immediately be replaced by a good car. It costs $1,000 per year to operate a good car and $1,500 to operate a fair car. Should I replace my car as soon as it becomes a fair car, or should I drive my car until it breaks down? Assume that the cost of operating a car during a year depends on the type of car on hand at the beginning of the year (after a new car, if any, arrives).

5 A square matrix is said to be doubly stochastic if its entries are all nonnegative and the entries in each row and each column sum to 1. For any ergodic, doubly stochastic matrix, show that all states have the same steady-state probability.

6 This problem will show why steady-state probabilities are sometimes referred to as stationary probabilities. Let $\pi_1, \pi_2, \ldots, \pi_s$ be the steady-state probabilities for an ergodic chain with transition matrix P. Also suppose that with probability π_i, the Markov chain begins in state i.

a What is the probability that after one transition, the system will be in state i? (*Hint:* Use Equation (8).)

b For any value of $n(n = 1, 2, \ldots)$, what is the probability that a Markov chain will be in state i after n transitions?

c Why are steady-state probabilities sometimes called stationary probabilities?

7 Consider two stocks. Stock 1 always sells for $10 or $20. If stock 1 is selling for $10 today, there is a .80 chance that it will sell for $10 tomorrow. If it is selling for $20 today, there is a .90 chance that it will sell for $20 tomorrow. Stock 2 always sells for $10 or $25. If stock 2 sells today for $10, there is a .90 chance that it will sell tomorrow for $10. If it sells today for $25, there is a .85 chance that it will sell tomorrow for $25. On the average, which stock will sell for a higher price? Find and interpret all mean first passage times.

8 Three balls are divided between two containers. During each period a ball is randomly chosen and switched to the other container.

a Find (in the steady state) the fraction of the time that a container will contain 0, 1, 2, or 3 balls.

b If container 1 contains no balls, on the average how many periods will go by before it again contains no balls? (*Note:* This is a special case of the Ehrenfest Diffusion model, which is used in biology to model diffusion through a membrane.)

9 Two types of squirrels—gray and black—have been seen in Pine Valley. At the beginning of each year, we determine which of the following is true:

There are only gray squirrels in Pine Valley.

There are only black squirrels in Pine Valley.

There are both gray and black squirrels in Pine Valley.

There are no squirrels in Pine Valley.

Over the course of many years, the following transition matrix has been estimated.

	Gray	Black	Both	Neither
Gray	.7	.2	.05	.05
Black	.2	.6	.1	.1
Both	.1	.1	.8	0
Neither	.05	.05	.1	.8

a During what fraction of years will gray squirrels be living in Pine Valley?

b During what fraction of years will black squirrels be living in Pine Valley?

Group B

10 Payoff Insurance Company charges a customer according to his or her accident history. A customer who has had no accident during the last two years is charged a $100 annual premium. Any customer who has had an accident during each of the last two years is charged a $400 annual premium. A customer who has had an accident during only one of the last two years is charged an annual premium of $300. A customer who has had an accident during the last year has a 10% chance of having an accident during the current year. If a customer has not had an accident during the last year, there is only a 3% chance that he or she will have an accident during the current year. During a given year, what is the average premium paid by a Payoff customer? (*Hint:* In case of difficulty, try a four-state Markov chain.)

11 Consider the following nonergodic chain:

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

a Why is the chain nonergodic?

b Explain why Theorem 1 fails for this chain. *Hint:* Find out if the following equation is true:

$$\lim_{n\to\infty} P_{12}(n) = \lim_{n\to\infty} P_{32}(n)$$

c Despite the fact that Theorem 1 fails, determine

$$\lim_{n\to\infty} P_{13}(n), \qquad \lim_{n\to\infty} P_{21}(n),$$

$$\lim_{n\to\infty} P_{43}(n), \qquad \lim_{n\to\infty} P_{41}(n)$$

12 Consider the following nonergodic chain:

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

a Why is this chain nonergodic?

b Explain why Theorem 1 fails for this chain. (*Hint:* Show that $\lim_{n\to\infty} P_{11}(n)$ does not exist by listing the pattern that $P_{11}(n)$ follows as n increases.)

13 An important machine is known to never last more than four months. During its first month of operation, it fails 10% of the time. If the machine completes its first month, then it fails during its second month 20% of the time. If the machine completes its second month of operation, then it will fail during its third month 50% of the time. If the machine completes its third month, then it is sure to fail by the end of the fourth month. At the beginning of each month, we must decide whether or not to replace our machine with a new machine. It costs $500 to purchase a new machine, but if a machine fails during a month, we incur a cost of $1,000 (due to factory downtime) and must replace the machine (at the beginning of the next month) with a new machine. Three maintenance policies are under consideration:

Policy 1 Plan to replace a machine at the beginning of its fourth month of operation.

Policy 2 Plan to replace a machine at the beginning of its third month of operation.

Policy 3 Plan to replace a machine at the beginning of its second month of operation.

Which policy will give the lowest average monthly cost?

14 Each month, customers are equally likely to demand 1 or 2 computers from a Pearco dealer. All orders must be met from current stock. Two ordering policies are under consideration:

Policy 1 If ending inventory is 2 units or less, order enough to bring next month's beginning inventory to 4 units.

Policy 2 If ending inventory is 1 unit or less, order enough to bring next month's beginning inventory up to 3 units.

The following costs are incurred by Pearco:

It costs $4,000 to order a computer.

It costs $100 to hold a computer in inventory for a month.

It costs $500 to place an order for computers. This is in addition to the per-customer cost of $4,000.

Which ordering policy has a lower expected monthly cost?

15 The Gotham City Maternity Ward contains 2 beds. Admissions are made only at the beginning of the day. Each day, there is a .5 probability that a potential admission will arrive. A patient can be admitted only if there is an open bed at the beginning of the day. Half of all patients are discharged after one day, and all patients that have stayed one day are discharged at the end of their second day.

a What is the fraction of days where all beds are utilized?

b On the average, what percentage of the beds are utilized?

17.6 Absorbing Chains

Many interesting applications of Markov chains involve chains in which some of the states are absorbing and the rest are transient states. Such a chain is called an **absorbing chain.** Consider an absorbing Markov chain: If we begin in a transient state, then eventually we are sure to leave the transient state and end up in one of the absorbing states. To see why we are interested in absorbing chains, we consider the following two absorbing chains.

EXAMPLE 7 Accounts Receivable

The accounts receivable situation of a firm is often modeled as an absorbing Markov chain.† Suppose a firm assumes that an account is uncollectable if the account is more than three months overdue. Then at the beginning of each month, each account may be classified into one of the following states:

State 1 New account

State 2 Payment on account is one month overdue.

State 3 Payment on account is two months overdue.

State 4 Payment on account is three months overdue.

†This example is based on Cyert, Davidson, and Thompson (1963).

State 5 Account has been paid.

State 6 Account is written off as bad debt.

Suppose that past data indicate that the following Markov chain describes how the status of an account changes from one month to the next month:

New 1 month 2 months 3 months Paid Bad debt

For example, if an account is two months overdue at the beginning of a month, there is a 40% chance that at the beginning of next month, the account will not be paid up (and therefore be three months overdue) and a 60% chance that the account will be paid up. To simplify our example, we assume that after three months, a debt is either collected or written off as a bad debt.

Once a debt is paid up or written off as a bad debt, the account is closed, and no further transitions occur. Hence, Paid and Bad Debt are absorbing states. Since every account will eventually be paid up or written off as a bad debt, New, 1 Month, 2 Months, and 3 Months are transient states. For example, a two-month overdue account can follow the path 2 Months–Collected, but there is no return path from Collected to 2 Months.

A typical new account will be absorbed as either a collected debt or a bad debt. A question of major interest is: What is the probability that a new account will eventually be collected? The answer is worked out later in this section.

EXAMPLE 8 Work-Force Planning

The law firm of Mason and Burger employs three types of lawyers: junior lawyers, senior lawyers, and partners. During a given year, there is a .15 probability that a junior lawyer will be promoted to senior lawyer and a .05 probability that he or she will leave the firm. Also, there is a .20 probability that a senior lawyer will be promoted to partner and a .10 probability that he or she will leave the firm. There is a .05 probability that a partner will leave the firm. The firm never demotes a lawyer.

There are many interesting questions the law firm might want to answer. For example, what is the probability that a newly hired junior lawyer will leave the firm before becoming a partner? On the average, how long does a newly hired junior lawyer stay with the firm? The answers are worked out later in this section.

We model the career path of a lawyer through Mason and Burger as an absorbing Markov chain with the following transition probability matrix:

Junior Senior Partner Leave as NP Leave as P

Junior	.80	.15	0	.05	0
Senior	0	.70	.20	.10	0
Partner	0	0	.95	0	.05
Leave as nonpartner	0	0	0	1	0
Leave as partner	0	0	0	0	1

The last two states are absorbing states, and all other states are transient. For example, Senior is a transient state, because there is a path from Senior to Leave as Nonpartner, but there is no path returning from Leave as Nonpartner to Senior (we assume that once a lawyer leaves the firm, he or she never returns).

For any absorbing chain, one might want to know certain things. (1) If the chain begins in a given transient state, and before we reach an absorbing state, what is the expected number of times that each state will be entered? How many periods do we expect to spend in a given transient state before absorption takes place? (2) If a chain begins in a given transient state, what is the probability that we end up in each absorbing state?

To answer these questions, we need to write the transition matrix with the states listed in the following order: transient states first, then absorbing states. For the sake of definiteness, let's assume that there are $s - m$ transient states ($t_1, t_2, \ldots, t_{s-m}$) and m absorbing states ($a_1, a_2, \ldots, a_m$). Then the transition matrix for the absorbing chain may be written as follows:

$$P = \begin{array}{r} \\ s-m \text{ rows} \\ m \text{ rows} \end{array} \begin{array}{c|c} s-m & m \\ \text{columns} & \text{columns} \\ \hline Q & R \\ \hline 0 & I \end{array}$$

In this format, the rows and column of P correspond (in order) to the states $t_1, t_2, \ldots, t_{s-m}, a_1, a_2, \ldots, a_m$. Here, I is an $m \times m$ identity matrix reflecting the fact that we can never leave an absorbing state: Q is an $(s - m) \times (s - m)$ matrix that represents transitions between transient states; R is an $(s - m) \times m$ matrix representing transitions from transient states to absorbing states; 0 is an $m \times (s - m)$ matrix consisting entirely of zeros. This reflects the fact that it is impossible to go from an absorbing state to a transient state.

Applying this notation to Example 7, we let

$$\begin{aligned} t_1 &= \text{New} \\ t_2 &= \text{1 Month} \\ t_3 &= \text{2 Months} \\ t_4 &= \text{3 Months} \\ a_1 &= \text{Paid} \\ a_2 &= \text{Bad Debt} \end{aligned}$$

Then for Example 7, the transition probability matrix may be written as

	New	1 month	2 months	3 months	Paid	Bad debt
New	0	.6	0	0	.4	0
1 month	0	0	.5	0	.5	0
2 months	0	0	0	.4	.6	0
3 months	0	0	0	0	.7	.3
Paid	0	0	0	0	1	0
Bad debt	0	0	0	0	0	1

Then $s = 6$, $m = 2$, and

$$Q = \begin{bmatrix} 0 & .6 & 0 & 0 \\ 0 & 0 & .5 & 0 \\ 0 & 0 & 0 & .4 \\ 0 & 0 & 0 & 0 \end{bmatrix}_{4\times 4} \qquad R = \begin{bmatrix} .4 & 0 \\ .5 & 0 \\ .6 & 0 \\ .7 & .3 \end{bmatrix}_{4\times 2}$$

For Example 8, we let

$$\begin{aligned} t_1 &= \text{Junior} \\ t_2 &= \text{Senior} \\ t_3 &= \text{Partner} \\ a_1 &= \text{Leave as nonpartner} \\ a_2 &= \text{Leave as partner} \end{aligned}$$

and we may write the transition probability matrix as

	Junior	Senior	Partner	Leave as NP	Leave as P
Junior	.80	.15	0	.05	0
Senior	0	.70	.20	.10	0
Partner	0	0	.95	0	.05
Leave as nonpartner	0	0	0	1	0
Leave as partner	0	0	0	0	1

Then $s = 5$, $m = 2$, and

$$Q = \begin{bmatrix} .80 & .15 & 0 \\ 0 & .70 & .20 \\ 0 & 0 & .95 \end{bmatrix}_{3\times 3} \qquad R = \begin{bmatrix} .05 & 0 \\ .10 & 0 \\ 0 & .05 \end{bmatrix}_{3\times 2}$$

We can now find out some facts about absorbing chains (see Kemeny and Snell (1960). (1) If the chain begins in a given transient state, and before we reach an absorbing state, what is the expected number of times that each state will be entered? How many periods do we expect to spend in a given transient state before absorption takes place? *Answer:* If we are at present in transient state t_i, the expected number of periods that will be spent in transient state t_j before absorption is the ijth element of the matrix $(I - Q)^{-1}$. (See Problem 12 at the end of this section for a proof.) (2) If a chain begins in a given transient state, what is the probability that we end up in each absorbing state? *Answer:* If we are at present in transient state t_i, the probability that we will eventually be absorbed in absorbing state a_j is the ijth element of the matrix $(I - Q)^{-1}R$. (See Problem 13 at the end of this section for a proof.)

The matrix $(I - Q)^{-1}$ is often referred to as the **Markov chain's fundamental matrix.** The reader interested in further study of absorbing chains is referred to Kemeny and Snell (1960).

EXAMPLE 7 **Accounts Receivable (Continued)**

1 What is the probability that a new account will eventually be collected?

2 What is the probability that a one-month-overdue account will eventually become a bad debt?

3 If the firm's sales average $100,000 per month, how much money per year will go uncollected?

Solution From our previous discussion, recall that

$$Q = \begin{bmatrix} 0 & .6 & 0 & 0 \\ 0 & 0 & .5 & 0 \\ 0 & 0 & 0 & .4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad R = \begin{bmatrix} .4 & 0 \\ .5 & 0 \\ .6 & 0 \\ .7 & .3 \end{bmatrix}$$

Then

$$I - Q = \begin{bmatrix} 1 & -.6 & 0 & 0 \\ 0 & 1 & -.5 & 0 \\ 0 & 0 & 1 & -.4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

By using the Gauss–Jordan method of Chapter 2, we find that

$$(I - Q)^{-1} = \begin{array}{c} \\ t_1 \\ t_2 \\ t_3 \\ t_4 \end{array} \begin{array}{c} \begin{array}{cccc} t_1 & t_2 & t_3 & t_4 \end{array} \\ \begin{bmatrix} 1 & .60 & .30 & .12 \\ 0 & 1 & .50 & .20 \\ 0 & 0 & 1 & .40 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

To answer questions 1–3, we need to compute

$$(I - Q)^{-1}R = \begin{array}{c} \\ t_1 \\ t_2 \\ t_3 \\ t_4 \end{array} \begin{array}{c} \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{bmatrix} .964 & .036 \\ .940 & .060 \\ .880 & .120 \\ .700 & .300 \end{bmatrix} \end{array}$$

Then

1 t_1 = New, a_1 = Paid. Thus, the probability that a new account is eventually collected is element 11 of $(I - Q)^{-1}R = .964$.

2 t_2 = 1 Month, a_2 = Bad Debt. Thus, the probability that a one-month overdue account turns into a bad debt is element 22 of $(I - Q)^{-1}R = .06$.

3 From answer 1, only 3.6% of all debts are uncollected. Since yearly accounts payable are $1,200,000, on the average, (.036)(1,200,000) = $43,200 per year will be uncollected.

EXAMPLE 8 **Work-Force Planning (Continued)**

1 What is the average length of time that a newly hired junior lawyer spends working for the firm?

2 What is the probability that a junior lawyer makes it to partner?

3 What is the average length of time that a partner spends with the firm (as a partner)?

Solution Recall that for Example 8,

$$Q = \begin{bmatrix} .80 & .15 & 0 \\ 0 & .70 & .20 \\ 0 & 0 & .95 \end{bmatrix} \qquad R = \begin{bmatrix} .05 & 0 \\ .10 & 0 \\ 0 & .05 \end{bmatrix}$$

Then

$$I - Q = \begin{bmatrix} .20 & -.15 & 0 \\ 0 & .30 & -.20 \\ 0 & 0 & .05 \end{bmatrix}$$

By using the Gauss–Jordan method of Chapter 2, we find that

IQinverse.xls

$$(I - Q)^{-1} = \begin{matrix} & t_1 & t_2 & t_3 \\ t_1 & 5 & 2.5 & 10 \\ t_2 & 0 & \frac{10}{3} & \frac{40}{3} \\ t_3 & 0 & 0 & 20 \end{matrix}$$

Then

	B	C	D	E	F
2					
3					
4		0.2	-0.15	0	
5	I-Q	0	0.3	-0.2	
6		0	0	0.05	
7					
8		5	2.5	10	
9	$(I-Q)^{-1}$	0	3.333333	13.33333	
10		0	0	20	
11					

FIGURE 8

$$(I - Q)^{-1}R = \begin{matrix} & a_1 & a_2 \\ t_1 & .50 & .50 \\ t_2 & \frac{1}{3} & \frac{2}{3} \\ t_3 & 0 & 1 \end{matrix}$$

Then

1 Expected time junior lawyer stays with firm = (expected time junior lawyer stays with firm as junior) + (expected time junior lawyer stays with firm as senior) + (expected time junior lawyer stays with firm as partner). Now

†Based on Bessent and Bessent (1980).

Expected time as junior = $(I - Q)_{11}^{-1} = 5$

Expected time as senior = $(I - Q)_{12}^{-1} = 2.5$

Expected time as partner = $(I - Q)_{13}^{-1} = 10$

Hence, the total expected time that a junior lawyer spends with the firm is $5 + 2.5 + 10 = 17.5$ years.

2 The probability that a new junior lawyer makes it to partner is just the probability that he or she leaves the firm as a partner. Since t_1 = Junior Lawyer and a_2 = Leave as Partner, the answer is element 12 of $(I - Q)^{-1}R = .50$.

3 Since t_3 = Partner, we seek the expected number of years that are spent in t_3, given that we begin in t_3. This is just element 33 of $(I - Q)^{-1} =$ 20 years. This is reasonable, because during each year, there is 1 chance in 20 that a partner will leave the firm, so it should take an average of 20 years before a partner leaves the firm.

†Based on Deming and Glasser (1968).
‡Based on Thompson and McNeal (1967).
§Based on Meredith (1973).

TABLE 4

	Accounting	Management Consulting	Division 1	Division 2	Division 3
Accounting	10%	30%	20%	20%	20%
Management	30%	20%	30%	0%	20%

REMARKS Computations with absorbing chains are greatly facilitated if we multiply matrices on a spreadsheet with the MMULT command and find the inverse of $(I - Q)$ with the MINVERSE function.

To use the Excel MINVERSE command to find $(I - Q)^{-1}$, we enter $(I - Q)$ into a spreadsheet (see cell range C4:E6 of file IQinverse.xls) and select the range (C8:E10) where we want to compute $(I - Q)^{-1}$. Next we type the formula

=MINVERSE(C4:E6)

in the upper left-hand corner (cell C8) of the output range C8:E10. Finally, we select **CONTROL SHIFT ENTER** (not just ENTER) to complete the computation of the desired inverse. The MINVERSE function must be entered with CONTROL SHIFT ENTER because it is an array function.

†This section covers topics that may be omitted with no loss of continuity.

We cannot edit or delete any part of a range computed by an array function. See Figure 8.

PROBLEMS

Group A

1[†] The State College admissions office has modeled the path of a student through State College as a Markov chain:

	F.	So.	J.	Sen.	Q.	G.
Freshman	.10	.80	0	0	.10	0
Sophmore	0	.10	.85	0	.05	0
Junior	0	0	.15	.80	.05	0
Senior	0	0	0	.10	.05	.85
Quits	0	0	0	0	1	0
Graduates	0	0	0	0	0	1

Each student's state is observed at the beginning of each fall semester. For example, if a student is a junior at the beginning of the current fall semester, there is an 80% chance that he will be a senior at the beginning of the next fall semester, a 15% chance that he will still be a junior, and a 5% chance that he will have quit. (We assume that once a student quits, he never reenrolls.)

a If a student enters State College as a freshman, how many years can he expect to spend as a student at State?

b What is the probability that a freshman graduates?

2[†] The *Herald Tribble* has obtained the following information about its subscribers: During the first year as subscribers, 20% of all subscribers cancel their subscriptions. Of those who have subscribed for one year, 10% cancel during the second year. Of those who have been subscribing for more than two years, 4% will cancel during any given year. On the average, how long does a subscriber subscribe to the *Herald Tribble?*

3 A forest consists of two types of trees: those that are 0–5 ft and those that are taller than 5 ft. Each year, 40% of all 0–5-ft tall trees die, 10% are sold for $20 each, 30% stay between 0 and 5 ft, and 20% grow to be more than 5 ft. Each year, 50% of all trees taller than 5 ft are sold for $50, 20% are sold for $30, and 30% remain in the forest.

a What is the probability that a 0–5-ft tall tree will die before being sold?

b If a tree (less than 5 ft) is planted, what is the expected revenue earned from that tree?

4[‡] Absorbing Markov chains are used in marketing to model the probability that a customer who is contacted by telephone will eventually buy a product. Consider a prospective customer who has never been called about purchasing a product. After one call, there is a 60% chance that the customer will express a low degree of interest in the product, a 30% chance of a high degree of interest, and a 10% chance the customer will be deleted from the company's list of prospective customers. Consider a customer who currently expresses a low degree of interest in the product. After another call, there is a 30% chance that the customer will purchase the product, a 20% chance the person will be deleted from the list, a 30% chance that the customer will still possess a low degree of interest, and a 20% chance that the customer will express a high degree of interest. Consider a customer who currently expresses a high degree of interest in the product. After another call, there is a 50% chance that the customer will have purchased the product, a 40% chance that the customer will still have a high degree of interest, and a 10% chance that the customer will have a low degree of interest.

a What is the probability that a new prospective customer will eventually purchase the product?

b What is the probability that a low-interest prospective customer will ever be deleted from the list?

c On the average, how many times will a new prospective customer be called before either purchasing the product or being deleted from the list?

5 Each week, the number of acceptable-quality units of a drug that are processed by a machine is observed: >100, 50–100, 1–50, 0 (indicating that the machine was broken during the week). Given last week's observation, the probability distribution of next week's observation is as follows.

	>100	50–100	1–50	0
>100	.8	.1	.05	.05
50–100	.1	.6	.1	.2
1–50	.1	.1	.5	.3
0	0	0	0	1

For example, if we observe a week in which more than 100 units are produced, then there is a .10 chance that during the next week 50–100 units are produced.

a Suppose last week the machine produced 200 units. On average, how many weeks will elapse before the machine breaks down?

b Suppose last week the machine produced 50 units. On average, how many weeks will elapse before the machine breaks down?

6 I now have \$2, and my goal is to have \$6. I will repeatedly flip a coin that has a .4 chance of coming up heads. If the coin comes up heads, I win the amount I bet. If the coin comes up tails, I lose the amount of my bet. Let us suppose I follow the **bold strategy** of betting Min(\$6 − current asset position, current asset position). This strategy (see Section 19.3) maximizes my chance of reaching my goal. What is the probability that I reach my goal?

7 Suppose I toss a fair coin, and the first toss comes up heads. If I keep tossing the coin until I either see two consecutive heads or two consecutive tails, what is the probability that I will see two consecutive heads before I see two consecutive tails?

8 Suppose each box of Corn Snaps cereal contains one of five different Harry Potter trading cards. On the average, how many boxes of cereal will I have to buy to obtain a complete set of trading cards?

Group B

9 In the gambler's ruin problem (Example 1), assume $p = .60$.

a What is the probability that I reach \$4?

b What is the probability that I am wiped out?

c What is the expected duration of the game?

10[§] In caring for elderly patients at a mental hospital, a major goal of the hospital is successful placement of the patients in boarding homes or nursing homes. The movement of patients between the hospital, outside homes, and the absorbing state (death) may be described by the following Markov chain (the unit of time is one month):

Hospital Homes Death

Each month that a patient spends in the hospital costs the state \$655, and each month that a patient spends in a home costs the state \$226. To improve the success rate of the placement of patients in homes, the state has recently begun a "geriatric resocialization program" (GRP) to prepare the patients for functioning in the homes. Some patients are placed in the GRP and then released to homes. These patients presumably are less likely to fail to adjust in the homes. Other patients continue to go directly from the hospital to homes without taking part in the GRP. The state pays \$680 for each month that a patient spends in the GRP. The movement of the patients through various states is governed by the following Markov chain:

	GRP	Hos.	Homes (GRP)	Homes (Direct)	Dead
GRP	.854	.028	.112	0	.006
Hospital	.013	.978	0	.003	.006
Homes (GRP)	.025	0	.969	0	.006
Homes (Direct)	0	.025	0	.969	.006
Dead	0	0	0	0	1

a Does the GRP save the state money?

b Under the old system and under the GRP, compute the expected number of months that a patient spends in the hospital.

11 Freezco, Inc., sells refrigerators. The company has issued a warranty on all refrigerators that requires free replacement of any refrigerator that fails before it is three years old. We are given the following information: (1) 3% of all new refrigerators fail during their first year of operation; (2) 5% of all one-year-old refrigerators fail during their second year of operation; and (3) 7% of all two-year-old refrigerators fail during their third year of operation. A replacement refrigerator is not covered by the warranty.

a Use Markov chain theory to predict the fraction of all refrigerators that Freezco will have to replace.

b Suppose that it costs Freezco \$500 to replace a refrigerator and that Freezco sells 10,000 refrigerators per year. If the company reduced the warranty period to two years, how much money in replacement costs would be saved?

12 For a Q matrix representing the transitions between transient states in an absorbing Markov chain, it can be shown that

$$(I - Q)^{-1} = I + Q + Q^2 + \cdots + Q^n + \cdots$$

a Explain why this expression for $(I - Q)^{-1}$ is plausible.

b Define m_{ij} = expected number of periods spent in transient state t_j before absorption, given that we begin in state t_i.

(Assume that the initial period is spent in state t_i.) Explain why m_{ij} = (probability that we are in state t_j initially) + (probability that we are in state t_j after first transition) + (probability that we are in state t_j after second transition) + · · · + (probability that we are in state t_j after nth transition) + · · ·.

c Explain why the probability that we are in state t_j initially = ijth entry of the $(s - m) \times (s - m)$ identity matrix. Explain why the probability that we are in state t_j after nth transition = ijth entry of Q^n.

d Now explain why $m_{ij} = ij$th entry of $(I - Q)^{-1}$.

13 Define

b_{ij} = probability of ending up in absorbing state a_j given that we begin in transient state t_i

r_{ij} = ijth entry of R

q_{ik} = ikth entry of Q

$B = (s - m) \times m$ matrix whose ijth entry is b_{ij}

Suppose we begin in state t_i. On our first transition, three types of events may happen:

Census.lng

Event 1 We go to absorbing state a_j (with probability r_{ij}).

Event 2 We go to an absorbing state other than a_j (with probability $\sum_{k \neq j} r_{ik}$).

Event 3 We go to transient state t_k (with probability q_{ik}).

a Explain why

$$b_{ij} = r_{ij} + \sum_{k=1}^{k=s-m} q_{ik} b_{kj}$$

b Now show that $b_{ij} = ij$th entry of $(R + QB)$ and that $B = R + QB$.

c Show that $B = (I - Q)^{-1}R$ and that $b_{ij} = ij$th entry of $B = (I - Q)^{-1}R$.

14 Consider an LP with five basic feasible solutions and a unique optimal solution. Assume that the simplex method begins at the worst basic feasible solution, and on each pivot the simplex is equally likely to move to any better basic feasible solution. On the average, how many pivots will be required to find the optimal solution to the LP?

Group C

15 General Motors has three auto divisions (1, 2, and 3). It also has an accounting division and a management consulting division. The question is: What fraction of the cost of the accounting and management consulting divisions should be allocated to each auto division? We assume that the entire cost of the accounting and management consulting departments must be allocated to the three auto divisions. During a given year, the work of the accounting division and management consulting division is allocated as shown in Table 4.

For example, accounting spends 10% of its time on problems generated by the accounting department, 20% of its time on work generated by division 3, and so forth. Each year, it costs \$63 million to run the accounting department and \$210 million to run the management consulting department. What fraction of these costs should be allocated to each auto division? Think of \$1 in costs incurred in accounting work. There is a .20 chance that this dollar should be allocated to each auto division, a .30 chance it should be allocated to consulting, and a .10 chance to accounting. If the dollar is allocated to an auto division, we know which division should be charged for that dollar. If the dollar is charged to consulting (for example), we repeat the process until the dollar is eventually charged to an auto division. Use knowledge of absorbing chains to figure out how to allocate the costs of running the accounting and management consulting departments among the three auto divisions.

16 A telephone sales force can model its contact with customers as a Markov chain. The six states of the chain are as follows:

State 1 Sale completed during most recent call

State 2 Sale lost during most recent call

State 3 New customer with no history

State 4 During most recent call, customer's interest level low

State 5 During most recent call, customer's interest level medium

State 6 During most recent call, customer's interest level high

Based on past phone calls, the following transition matrix has been estimated:

	1	2	3	4	5	6
1	1	0	0	0	0	0
2	0	1	0	0	0	0
3	.10	.30	0	.25	.20	.15
4	.05	.45	0	.20	.20	.10
5	.15	.10	0	.15	.25	.35
6	.20	.05	0	.15	.30	.30

a For a new customer, determine the average number of calls made before the customer buys the product or the sale is lost.

b What fraction of new customers will buy the product?

c What fraction of customers currently having a low degree of interest will buy the product?

d Suppose a call costs \$15 and a sale earns \$190 in revenue. Determine the "value" of each type of customer.

17 Seas Beginning sells clothing by mail order. An important question is: When should the company strike a customer from its mailing list? At present, the company does so if a customer fails to order from six consecutive catalogs. Management wants to know if striking a customer after failure to order from four consecutive catalogs will result in a higher profit per customer.

The following data are available: Six percent of all customers who receive a catalog for the first time place an order. If a customer placed an order from the last-received catalog, then there is a 20% chance he or she will order from the next catalog. If a customer last placed an order one catalog ago, there is a 16% chance he or she will order from the next catalog received. If a customer last placed an order two catalogs ago, there is a 12% chance he or she will place an order from the next catalog received. If a customer last placed an order three catalogs ago, there is an 8% chance he or she will place an order from the next catalog received. If a customer last placed an order four catalogs ago, there is a 4% chance he or she will place an order from the next catalog received. If a customer last placed an order five catalogs ago, there is a 2% chance he or she will place an order from the next catalog received.

It costs \$1 to send a catalog, and the average profit per order is \$15. To maximize expected profit per customer, should Seas Beginning cancel customers after six nonorders or four nonorders?

Hint: Model each customer's evolution as a Markov chain with possible states New, 0, 1, 2, 3, 4, 5, Canceled. A customer's state represents the number of catalogs received since the customer last placed an order. "New" means the customer received a catalog for the first time. "Canceled" means that the customer has failed to order from six consecutive catalogs. For example, suppose a customer placed the following sequence of orders (O) and nonorders (NO):

NO NO O NO NO O O NO NO O NO NO NO NO NO NO Canceled

Here we are assuming a customer is stricken from the mailing list after six consecutive nonorders. For this sequence of orders and nonorders, the states are (*i*th listed state occurs right before *i*th catalog is received)

New 1 2 0 1 2 0 0 1 2 0 1 2 3 4 5 Canceled

You should be able to figure (for each cancellation policy) the expected number of orders a customer will place before cancellation and the expected number of catalogs a customer will receive before cancellation. This will enable you to compute expected profit per customer.

TABLE 5

	Age of Blood (beginning of day)				
Chance of transfusion	0	1	2	3	4
Policy 1	.10	.20	.30	.40	.50
Policy 2	.50	.40	.30	.20	.10

†Based on Pegels and Jelmert (1970).

17.7 Work-Force Planning Models‡

Many organizations, like the Mason and Burger law firm of Example 8, employ several categories of workers. For long-term planning purposes, it is often useful to be able to predict the number of employees of each type who will (if present trends continue) be available in the steady state. Such predictions can be made via an analysis similar to the one in Section 17.5 of steady-state probabilities for Markov chains.

More formally, consider an organization whose members are classified at any point in time into one of s groups (labeled $1, 2, \ldots, s$). During

TABLE 6

Age	Death Probability
0	0.007557
1–4	0.000383
5–9	0.000217
10–14	0.000896
15–24	0.001267
25–34	0.002213
35–44	0.004459
45–54	0.010941
55–64	0.025384
65–84	0.058031
85+	0.15327

‡Based on Babich (1992).

every time period, a fraction p_{ij} of those who begin a time period in group i begin the next time period in group j. Also, during every time period, a fraction $p_{i,s+1}$ of all group i members leave the organization. Let P be the $s \times$

SUMMARY

Let $\mathbf{X}_t$ be the value of a system's characteristic at time t. A **discrete-time stochastic process** is simply a description of the relation between the random variables $\mathbf{X}_0$, $\mathbf{X}_1$, $\mathbf{X}_2$, ... A discrete-time stochastic process is a **Markov chain** if, for $t = 0, 1, 2, \ldots$ and all states,

$$P(\mathbf{X}_{t+1} = i_{t+1}|\mathbf{X}_t = i_t, \mathbf{X}_{t-1} = i_{t-1}, \ldots, \mathbf{X}_1 = i_1, \mathbf{X}_0 = i_0) = P(\mathbf{X}_{t+1} = i_{t+1}|\mathbf{X}_t = i_t)$$

For a stationary Markov chain, the **transition probability** p_{ij} is the probability that given the system is in state i at time t, the system will be in state j at time $t + 1$.

The vector $\mathbf{q} = [q_1 \quad q_2 \quad \cdots \quad q_s]$ is the **initial probability distribution** for the Markov chain. $P(\mathbf{X}_0 = i)$ is given by q_i.

n-Step Transition Probabilities

The ***n*-step transition probability,** $p_{ij}(n)$, is the probability that n periods from now, the state will be j, given that the current state is i. $P_{ij}(n) = ij$th element of P^n.

Given the intial probability vector $\mathbf{q}$, the probability of being in state j at time n is given by $\mathbf{q}$(column j of P^n).

Classification of States in a Markov Chain

Given two states i and j, a **path** from i to j is a sequence of transitions that begins in i and ends in j, such that each transition in the sequence has a positive probability of occurring. A state j is **reachable** from a state i if there is a path leading from i to j. Two states i and j are said to **communicate** if j is reachable from i, and i is reachable from j.

A set of states S in a Markov chain is a **closed set** if no state outside of S is reachable from any state in S.

A state i is an **absorbing state** if $p_{ii} = 1$. A state i is a **transient state** if there exists a state j that is reachable from i, but the state i is not reachable from state j.

If a state is not transient, it is a **recurrent state.** A state i is **periodic** with period $k > 1$ if all paths leading from state i back to state i have a length that is a multiple of k. If a recurrent state is not periodic, it is **aperiodic.** If all states in a chain are recurrent, aperiodic, and communicate with each other, the chain is said to be **ergodic.**

Steady-State Probabilities

Let P be the transition probability matrix for an ergodic Markov chain with states 1, 2, . . . , s (with ijth element p_{ij}). After a large number of periods have elapsed, the probability (call it π_j) that the Markov chain is in state j is independent of the initial state. The long-run, or **steady-state,** probability π_j may be found by solving the following set of linear equations:

$$\pi_j = \sum_{k=1}^{k=s} \pi_k p_{kj} \qquad (j = 1, 2, \ldots, s;\ \text{omit one of these equations})$$

$$\pi_1 + \pi_2 + \cdots + \pi_s = 1$$

Absorbing Chains

A Markov chain in which one or more states is an absorbing state is an **absorbing Markov chain.** To answer important questions about an absorbing Markov chain, we list the states in the following order: transient states first, then absorbing states. Assume there are $s - m$ transient states ($t_1, t_2, \ldots, t_{s-m}$) and m absorbing states ($a_1, a_2, \ldots, a_m$). Write the transition probability matrix P as follows:

$$P = \begin{array}{r} \\ s-m \text{ rows} \\ m \text{ rows} \end{array} \begin{array}{c} \begin{array}{cc} s-m \text{ columns} & m \text{ columns} \end{array} \\ \left[\begin{array}{c|c} Q & R \\ \hline 0 & I \end{array}\right] \end{array}$$

The following questions may now be answered. (1) If the chain begins in a given transient state, and before we reach an absorbing state, what is the expected number of times that each state will be entered? How many periods do we expect to spend in a given transient state before absorption takes place? *Answer:* If we are at present in transient state t_i, the expected number of periods that will be spent in transient state t_j before absorption is the ijth element of the matrix $(I - Q)^{-1}$. (2) If a chain begins in a given transient state, what is the probability that we will end up in each absorbing state? *Answer:* If we are at present in transient state t_i, the probability that we will eventually be absorbed in absorbing state a_j is the ijth element of the matrix $(I - Q)^{-1}R$.

Work-Force Planning Models

For an organization in which each member is classified into one of s groups,

p_{ij} = fraction of members beginning a time period in group i who begin the next time period in group j

$p_{i,s+1}$ = fraction of all group i members who leave the organization during a period

P = $s \times (s + 1)$ matrix whose ijth entry is p_{ij}

H_i = number of group i members hired at the beginning of each period

N_i = limiting number (if it exists) of group i members

N_i may be found by equating the number of people per period who enter group i with the number of people per period who leave group i. Thus, ($N_1, N_2, \ldots, N_s$) may be found by solving

$$H_i + \sum_{k \neq i} N_k p_{ki} = N_i \sum_{k \neq i} p_{ik} \qquad (i = 1, 2, \ldots, s)$$

$$\begin{aligned} H_1 &= (.15 + .05)50 && \text{(Junior lawyers)} \\ (.15)50 + H_2 &= (.20 + .10)30 && \text{(Senior lawyers)} \\ (.20)30 + H_3 &= (.05)10 && \text{(Partners)} \end{aligned}$$

The unique solution to this system of equations is $H_1 = 10, H_2 = 1.5, H_3 = -5.5$. This means that to maintain the desired steady-state census, Mason and Burger would have to fire 5.5 partners each year. This is reasonable, because an average of $.20(30) = 6$ senior lawyers become partners every year, and once a senior lawyer becomes a partner, he or she stays a partner for an average of 20 years. This shows that to keep the number of partners down to 10, several partners must be released each year. An alternative solution might be to reduce (below its current value of .20) the fraction of senior lawyers who become partners during each year.

[†]Based on Flamholtz, Geis, and Perle (1984).

†Based on Flamholtz, Geis, and Perle (1984).

‡Based on Eppen and Fama (1970).

[†]Based on Liu, Wang, and Guh (1991).
[‡]Based on "Statisticians Count Euros and Find More Than Money," *New York Times,* July 2, 2002.