BIJU PATNAIK UNIVERSITY OF TECHNOLOGY, ODISHA

Lecture Notes

On

INTRODUCTION TO OPERATION RESEARCH

Prepared by,

Dr. Subhendu Kumar Rath,

BPUT, Odisha.

The Markov chains to be considered in this chapter have the following properties:

1. A finite number of states.
2. Stationary transition probabilities.

We also will assume that we know the initial probabilities $P\{X_0 \quad i\}$ for all i.

Formulating the Inventory Example as a Markov Chain

Returning to the inventory example developed in the preceding section, recall that X_t is the number of cameras in stock at the end of week t (before ordering any more), where X_t represents the *state of the system* at time t. Given that the current state is $X_t \quad i$, the expression at the end of Sec. 16.1 indicates that $X_{t \quad 1}$ depends only on $D_{t \quad 1}$ (the demand in week $t \quad 1$) and X_t. Since $X_{t \quad 1}$ is independent of any past history of the inventory system, the stochastic process $\{X_t\}$ ($t \quad 0, 1, \ldots$) has the *Markovian property* and so is a Markov chain.

Now consider how to obtain the (one-step) transition probabilities, i.e., the elements of the (one-step) *transition matrix*

$$\mathbf{P} \quad \begin{array}{c} \text{State} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} p_{00} & p_{01} & p_{02} & p_{03} \\ p_{10} & p_{11} & p_{12} & p_{13} \\ p_{20} & p_{21} & p_{22} & p_{23} \\ p_{30} & p_{31} & p_{32} & p_{33} \end{bmatrix} \end{array}$$

given that $D_{t \quad 1}$ has a Poisson distribution with a mean of 1. Thus,

$$P\{D_{t \quad 1} \quad n\} \quad \frac{(1)^n e^{\quad 1}}{n!}, \qquad \text{for } n \quad 0, 1, \ldots,$$

so

$$P\{D_{t \quad 1} \quad 0\} \quad e^{\quad 1} \quad 0.368,$$
$$P\{D_{t \quad 1} \quad 1\} \quad e^{\quad 1} \quad 0.368,$$
$$P\{D_{t \quad 1} \quad 2\} \quad \frac{1}{2}e^{\quad 1} \quad 0.184,$$
$$P\{D_{t \quad 1} \quad 3\} \quad 1 \quad P\{D_{t \quad 1} \quad 2\} \quad 1 \quad (0.368 \quad 0.368 \quad 0.184) \quad 0.080.$$

For the first row of **P**, we are dealing with a transition from state $X_t \quad 0$ to some state $X_{t \quad 1}$. As indicated at the end of Sec. 16.1,

$$X_{t \quad 1} \quad \max\{3 \quad D_{t \quad 1}, 0\} \qquad \text{if } X_t \quad 0.$$

Therefore, for the transition to $X_{t \quad 1} \quad 3$ or $X_{t \quad 1} \quad 2$ or $X_{t \quad 1} \quad 1$,

$$p_{03} \quad P\{D_{t \quad 1} \quad 0\} \quad 0.368,$$
$$p_{02} \quad P\{D_{t \quad 1} \quad 1\} \quad 0.368,$$
$$p_{01} \quad P\{D_{t \quad 1} \quad 2\} \quad 0.184.$$

A transition from X_t 0 to $X_{t\ \ 1}$ 0 implies that the demand for cameras in week t 1 is 3 or more after 3 cameras are added to the depleted inventory at the beginning of the week, so

$$p_{00} \quad P\{D_{t\ \ 1} \quad 3\} \quad 0.080.$$

For the other rows of **P**, the formula at the end of Sec. 16.1 for the next state is

$$X_{t\ \ 1} \quad \max\{X_t \quad D_{t\ \ 1}, 0\} \qquad \text{if } X_{t\ \ 1} \quad 1.$$

This implies that $X_{t\ \ 1}$ X_t, so p_{12} 0, p_{13} 0, and p_{23} 0. For the other transitions,

$$p_{11} \quad P\{D_{t\ \ 1} \quad 0\} \quad 0.368,$$
$$p_{10} \quad P\{D_{t\ \ 1} \quad 1) \quad 1 \quad P\{D_{t\ \ 1} \quad 0\} \quad 0.632,$$
$$p_{22} \quad P\{D_{t\ \ 1} \quad 0\} \quad 0.368,$$
$$p_{21} \quad P\{D_{t\ \ 1} \quad 1\} \quad 0.368,$$
$$p_{20} \quad P\{D_{t\ \ 1} \quad 2\} \quad 1 \quad P\{D_{t\ \ 1} \quad 1\} \quad 1 \quad (0.368 \quad 0.368) \quad 0.264.$$

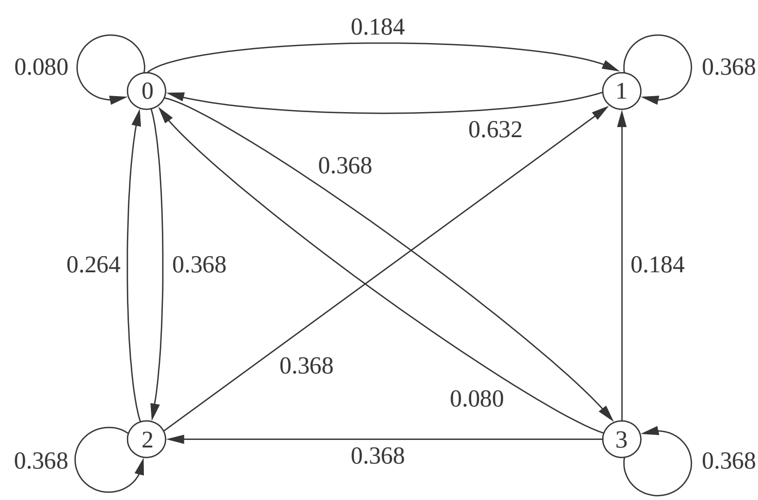
For the last row of **P**, week t 1 begins with 3 cameras in inventory, so the calculations for the transition probabilities are exactly the same as for the first row. Consequently, the complete transition matrix is

P

State	0	1	2	3
0	0.080	0.184	0.368	0.368
1	0.632	0.368	0	0
2	0.264	0.368	0.368	0
3	0.080	0.184	0.368	0.368

The information given by this transition matrix can also be depicted graphically with the state transition diagram in Fig. 16.1. The four possible states for the number of cameras on hand at the end of a week are represented by the four nodes (circles) in the diagram. The

FIGURE 16.1
State transition diagram for the inventory example for a camera store.



arrows show the possible transitions from one state to another, or sometimes from a state back to itself, when the camera store goes from the end of one week to the end of the next week. The number next to each arrow gives the probability of that particular transition occurring next when the camera store is in the state at the base of the arrow.

Additional Examples of Markov Chains

A Stock Example. Consider the following model for the value of a stock. At the end of a given day, the price is recorded. If the stock has gone up, the probability that it will go up tomorrow is 0.7. If the stock has gone down, the probability that it will go up tomorrow is only 0.5. This is a Markov chain, where state 0 represents the stock's going up and state 1 represents the stock's going down. The transition matrix is given by

	State	0	1
P	0	0.7	0.3
	1	0.5	0.5

A Second Stock Example. Suppose now that the stock market model is changed so that the stock's going up tomorrow depends upon whether it increased today *and* yesterday. In particular, if the stock has increased for the past two days, it will increase tomorrow with probability 0.9. If the stock increased today but decreased yesterday, then it will increase tomorrow with probability 0.6. If the stock decreased today but increased yesterday, then it will increase tomorrow with probability 0.5. Finally, if the stock decreased for the past two days, then it will increase tomorrow with probability 0.3. If we define the state as representing whether the stock goes up or down today, the system is no longer a Markov chain. However, we can transform the system to a Markov chain by defining the states as follows:[1]

State 0: The stock increased both today and yesterday.
State 1: The stock increased today and decreased yesterday.
State 2: The stock decreased today and increased yesterday.
State 3: The stock decreased both today and yesterday.

This leads to a four-state Markov chain with the following transition matrix:

	State	0	1	2	3
P	0	0.9	0	0.1	0
	1	0.6	0	0.4	0
	2	0	0.5	0	0.5
	3	0	0.3	0	0.7

A Gambling Example. Another example involves gambling. Suppose that a player has \$1 and with each play of the game wins \$1 with probability p 0 or loses \$1 with probability 1 p. The game ends when the player either accumulates \$3 or goes broke.

[1]This example demonstrates that Markov chains are able to incorporate arbitrary amounts of history, but at the cost of significantly increasing the number of states.

This game is a Markov chain with the states representing the player's current holding of money, that is, 0, \$1, \$2, or \$3, and with the transition matrix given by

$$\mathbf{P} = \begin{array}{c} \text{State} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1-p & 0 & p & 0 \\ 0 & 1-p & 0 & p \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Note that in both the inventory and gambling examples, the numeric labeling of the states that the process reaches coincides with the physical expression of the system—i.e., actual inventory levels and the player's holding of money, respectively—whereas the numeric labeling of the states in the stock examples has no physical significance.

16.3 CHAPMAN-KOLMOGOROV EQUATIONS

Section 16.2 introduced the n-step transition probability $p_{ij}^{(n)}$. The following *Chapman-Kolmogorov equations* provide a method for computing these n-step transition probabilities:

$$p_{ij}^{(n)} = \sum_{k=0}^{M} p_{ik}^{(m)} p_{kj}^{(n-m)}, \qquad \begin{aligned} &\text{for all } i = 0, 1, \ldots, M, \\ &j = 0, 1, \ldots, M, \\ &\text{and any } m = 1, 2, \ldots, n-1, \\ &n = m+1, m+2, \ldots.^{1} \end{aligned}$$

These equations point out that in going from state i to state j in n steps, the process will be in some state k after exactly m (less than n) states. Thus, $p_{ik}^{(m)} p_{kj}^{(n-m)}$ is just the conditional probability that, given a starting point of state i, the process goes to state k after m steps and then to state j in $n - m$ steps. Therefore, summing these conditional probabilities over all possible k must yield $p_{ij}^{(n)}$. The special cases of $m = 1$ and $m = n - 1$ lead to the expressions

$$p_{ij}^{(n)} = \sum_{k=0}^{M} p_{ik} p_{kj}^{(n-1)}$$

and

$$p_{ij}^{(n)} = \sum_{k=0}^{M} p_{ik}^{(n-1)} p_{kj},$$

for all states i and j. These expressions enable the n-step transition probabilities to be obtained from the one-step transition probabilities recursively. This recursive relationship is best explained in matrix notation (see Appendix 4). For $n = 2$, these expressions become

$$p_{ij}^{(2)} = \sum_{k=0}^{M} p_{ik} p_{kj}, \qquad \text{for all states } i \text{ and } j,$$

[1]These equations also hold in a trivial sense when $m = 0$ or $m = n$, but $m = 1, 2, \ldots, n - 1$ are the only interesting cases.

where the $p_{ij}^{(2)}$ are the elements of a matrix $\mathbf{P}^{(2)}$. Also note that these elements are obtained by multiplying the matrix of one-step transition probabilities by itself; i.e.,

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

In the same manner, the above expressions for $p_{ij}^{(n)}$ when $m = 1$ and $m = n - 1$ indicate that the matrix of n-step transition probabilities is

$$\begin{aligned}\mathbf{P}^{(n)} &= \mathbf{P}\mathbf{P}^{(n-1)} = \mathbf{P}^{(n-1)}\mathbf{P}\\ &= \mathbf{P}\mathbf{P}^{n-1} = \mathbf{P}^{n-1}\mathbf{P}\\ &= \mathbf{P}^n.\end{aligned}$$

Thus, the n-step transition probability matrix $\mathbf{P}^n$ can be obtained by computing the nth power of the one-step transition matrix $\mathbf{P}$.

n-Step Transition Matrices for the Inventory Example

Returning to the inventory example, its one-step transition matrix $\mathbf{P}$ obtained in Sec. 16.2 can now be used to calculate the two-step transition matrix $\mathbf{P}^{(2)}$ as follows:

$$\mathbf{P}^{(2)} = \mathbf{P}^2 = \begin{bmatrix} 0.080 & 0.184 & 0.368 & 0.368 \\ 0.632 & 0.368 & 0 & 0 \\ 0.264 & 0.368 & 0.368 & 0 \\ 0.080 & 0.184 & 0.368 & 0.368 \end{bmatrix} \begin{bmatrix} 0.080 & 0.184 & 0.368 & 0.368 \\ 0.632 & 0.368 & 0 & 0 \\ 0.264 & 0.368 & 0.368 & 0 \\ 0.080 & 0.184 & 0.368 & 0.368 \end{bmatrix}$$

$$= \begin{bmatrix} 0.249 & 0.286 & 0.300 & 0.165 \\ 0.283 & 0.252 & 0.233 & 0.233 \\ 0.351 & 0.319 & 0.233 & 0.097 \\ 0.249 & 0.286 & 0.300 & 0.165 \end{bmatrix}.$$

For example, given that there is one camera left in stock at the end of a week, the probability is 0.283 that there will be no cameras in stock 2 weeks later, that is, $p_{10}^{(2)} = 0.283$. Similarly, given that there are two cameras left in stock at the end of a week, the probability is 0.097 that there will be three cameras in stock 2 weeks later, that is, $p_{23}^{(2)} = 0.097$.

The four-step transition matrix can also be obtained as follows:

$$\mathbf{P}^{(4)} = \mathbf{P}^4 = \mathbf{P}^{(2)} \cdot \mathbf{P}^{(2)}$$

$$= \begin{bmatrix} 0.249 & 0.286 & 0.300 & 0.165 \\ 0.283 & 0.252 & 0.233 & 0.233 \\ 0.351 & 0.319 & 0.233 & 0.097 \\ 0.249 & 0.286 & 0.300 & 0.165 \end{bmatrix} \begin{bmatrix} 0.249 & 0.286 & 0.300 & 0.165 \\ 0.283 & 0.252 & 0.233 & 0.233 \\ 0.351 & 0.319 & 0.233 & 0.097 \\ 0.249 & 0.286 & 0.300 & 0.165 \end{bmatrix}$$

$$= \begin{bmatrix} 0.289 & 0.286 & 0.261 & 0.164 \\ 0.282 & 0.285 & 0.268 & 0.166 \\ 0.284 & 0.283 & 0.263 & 0.171 \\ 0.289 & 0.286 & 0.261 & 0.164 \end{bmatrix}.$$

For example, given that there is one camera left in stock at the end of a week, the probability is 0.282 that there will be no cameras in stock 4 weeks later, that is, $p_{10}^{(4)} = 0.282$.

Similarly, given that there are two cameras left in stock at the end of a week, the probability is 0.171 there will be three cameras in stock 4 weeks later, that is, $p_{23}^{(4)} \quad 0.171$.

Your OR Courseware includes a routine for calculating $\mathbf{P}^{(n)} \quad \mathbf{P}^n$ for any positive integer $n \quad 99$.

Unconditional State Probabilities

Recall that one- or n-step transition probabilities are *conditional* probabilities; for example, $P\{X_n \quad j \mid X_0 \quad i\} \quad p_{ij}^{(n)}$. If the *unconditional* probability $P\{X_n \quad j\}$ is desired, it is necessary to specify the probability distribution of the initial state, namely, $P\{X_0 \quad i\}$ for $i \quad 0, 1, \ldots, M$. Then

$$P\{X_n \quad j\} \quad P\{X_0 \quad 0\}\, p_{0j}^{(n)} \quad P\{X_0 \quad 1\} p_{1j}^{(n)} \qquad\qquad P\{X_0 \quad M\} p_{Mj}^{(n)}.$$

In the inventory example, it was assumed that initially there were 3 units in stock, that is, $X_0 \quad 3$. Thus, $P\{X_0 \quad 0\} \quad P\{X_0 \quad 1\} \quad P\{X_0 \quad 2\} \quad 0$ and $P\{X_0 \quad 3\} \quad 1$. Hence, the (unconditional) probability that there will be three cameras in stock 2 weeks after the inventory system began is $P\{X_2 \quad 3\} \quad (1)p_{33}^{(2)} \quad 0.165$.

16.4 CLASSIFICATION OF STATES OF A MARKOV CHAIN

It is evident that the transition probabilities associated with the states play an important role in the study of Markov chains. To further describe the properties of Markov chains, it is necessary to present some concepts and definitions concerning these states.

State j is said to be **accessible** from state i if $p_{ij}^{(n)} \quad 0$ for some $n \quad 0$. (Recall that $p_{ij}^{(n)}$ is just the conditional probability of being in state j after n steps, starting in state i.) Thus, state j being accessible from state i means that it is possible for the system to enter state j eventually when it starts from state i. In the inventory example, $p_{ij}^{(2)} \quad 0$ for all i and j, so every state is accessible from every other state. In general, a sufficient condition for *all* states to be accessible is that there exists a value of n for which $p_{ij}^{(n)} \quad 0$ for all i and j.

In the gambling example given at the end of Sec. 16.2, state 2 is not accessible from state 3. This can be deduced from the context of the game (once the player reaches state 3, the player never leaves this state), which implies that $p_{32}^{(n)} \quad 0$ for all $n \quad 0$. However, even though state 2 is *not* accessible from state 3, state 3 *is* accessible from state 2 since, for $n \quad 1$, the transition matrix given at the end of Sec. 16.2 indicates that $p_{23} \quad p \quad 0$.

If state j is accessible from state i and state i is accessible from state j, then states i and j are said to **communicate.** In the inventory example, all states communicate. In the gambling example, states 2 and 3 do not. In general,

1. Any state communicates with itself (because $p_{ii}^{(0)} \quad P\{X_0 \quad i \mid X_0 \quad i\} \quad 1$).
2. If state i communicates with state j, then state j communicates with state i.
3. If state i communicates with state j and state j communicates with state k, then state i communicates with state k.

Properties 1 and 2 follow from the definition of states communicating, whereas property 3 follows from the Chapman-Kolmogorov equations.

As a result of these three properties of communication, the states may be partitioned into one or more separate **classes** such that those states that communicate with each other are in the same class. (A class may consist of a single state). If there is only one class, i.e., all the states communicate, the Markov chain is said to be **irreducible.** In the inventory example, the Markov chain is irreducible. In the first stock example in Sec. 16.2, the Markov chain is irreducible. The gambling example contains three classes. State 0 forms a class, state 3 forms a class, and states 1 and 2 form a class.

Recurrent States and Transient States

It is often useful to talk about whether a process entering a state will ever return to this state. Here is one possibility.

> A state is said to be a **transient** state if, upon entering this state, the process *may never return* to this state again. Therefore, state i is transient if and only if there exists a state j (j i) that is accessible from state i but not vice versa, that is, state i is not accessible from state j.

Thus, if state i is transient and the process visits this state, there is a positive probability (perhaps even a probability of 1) that the process will later move to state j and so will never return to state i. Consequently, a transient state will be visited only a finite number of times.

When starting in state i, another possibility is that the process *definitely* will return to this state.

> A state is said to be a **recurrent** state if, upon entering this state, the process *definitely will return* to this state again. Therefore, a state is recurrent if and only if it is not transient.

Since a recurrent state definitely will be revisited after each visit, it will be visited infinitely often if the process continues forever.

If the process enters a certain state and then stays in this state at the next step, this is considered a *return* to this state. Hence, the following kind of state is a special type of recurrent state.

> A state is said to be an **absorbing** state if, upon entering this state, the process *never will leave* this state again. Therefore, state i is an absorbing state if and only if p_{ii} 1.

We will discuss absorbing states further in Sec. 16.7.

Recurrence is a class property. That is, all states in a class are either recurrent or transient. Furthermore, in a finite-state Markov chain, not all states can be transient. Therefore, all states in an irreducible finite-state Markov chain are recurrent. Indeed, one can identify an irreducible finite-state Markov chain (and therefore conclude that all states are recurrent) by showing that all states of the process communicate. It has already been pointed out that a sufficient condition for *all* states to be accessible (and therefore communicate with each other) is that there exists a value of n for which $p_{ij}^{(n)}$ 0 for all i and j. Thus, all states in the inventory example are recurrent, since $p_{ij}^{(2)}$ is positive for all i and j. Similarly, the first stock example contains only recurrent states, since p_{ij} is positive for all i and j. By calculating $p_{ij}^{(2)}$ for all i and j in the second stock example in Sec. 16.2, it follows that all states are recurrent since $p_{ij}^{(2)}$ 0.

As another example, suppose that a Markov chain has the following transition matrix:

$$\mathbf{P} = \begin{array}{c} \text{State} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Note that state 2 is an absorbing state (and hence a recurrent state) because if the process enters state 2 (row 3 of the matrix), it will never leave. State 3 is a transient state because if the process is in state 3, there is a positive probability that it will never return. The probability is $\frac{1}{3}$ that the process will go from state 3 to state 2 on the first step. Once the process is in state 2, it remains in state 2. State 4 also is a transient state because if the process starts in state 4, it immediately leaves and can never return. States 0 and 1 are recurrent states. To see this, observe from **P** that if the process starts in either of these states, it can never leave these two states. Furthermore, whenever the process moves from one of these states to the other one, it always will return to the original state eventually.

Periodicity Properties

Another useful property of Markov chains is *periodicities*. The **period** of state i is defined to be the integer t (t 1) such that $p_{ii}^{(n)}$ 0 for all values of n other than t, $2t$, $3t$, . . . and t is the largest integer with this property. In the gambling example (end of Section 16.2), starting in state 1, it is possible for the process to enter state 1 only at times 2, 4, . . . , so state 1 has period 2. The reason is that the player can break even (be neither winning nor losing) only at times 2, 4, . . . , which can be verified by calculating $p_{11}^{(n)}$ for all n and noting that $p_{11}^{(n)}$ 0 for n odd.

If there are two consecutive numbers s and s 1 such that the process can be in state i at times s and s 1, the state is said to have period 1 and is called an **aperiodic** state.

Just as recurrence is a class property, it can be shown that periodicity is a class property. That is, if state i in a class has period t, the all states in that class have period t. In the gambling example, state 2 also has period 2 because it is in the same class as state 1 and we noted above that state 1 has period 2.

In a finite-state Markov chain, recurrent states that are aperiodic are called **ergodic** states. A Markov chain is said to be *ergodic* if all its states are ergodic states.

16.5 LONG-RUN PROPERTIES OF MARKOV CHAINS

Steady-State Probabilities

In Sec. 16.3 the four-step transition matrix for the inventory example was obtained. It will now be instructive to examine the eight-step transition probabilities given by the matrix

$\mathbf{P}^{(8)}$ $\mathbf{P}^8$ $\mathbf{P}^4$ $\mathbf{P}^4$

State	3	0		1
			0.286	
0		0.286	0.286	0.285
1				0.285
2		0.286		0.285

0.285

2 3

0.264 0.264 0.264 0.264

0.166

0.166

0.166

0.166

Notice the rather remarkable fact that each of the four rows has identical entries. This implies that the probability of being in state j after 8 weeks is essentially independent of the initial level of inventory. In other words, it appears that there is a limiting probability that the system will be in each state j after a large number of transitions, and that this probability is independent of the initial state. These properties of the long-run behavior of finite-state Markov chains do, in fact, hold under relatively general conditions, as summarized below.

For any irreducible ergodic Markov chain, $\lim_{n\to} p_{ij}^{(n)}$ exists and is independent of i. Furthermore,

$$\lim_{n\to} p_{ij}^{(n)} \quad {}_j \quad 0,$$

where the $_j$ uniquely satisfy the following **steady-state equations**

$$_j \quad \overset{M}{\underset{i \quad 0}{}} \quad {}_i p_{ij}, \qquad \text{for } j \quad 0, 1, \ldots, M,$$

$$\overset{M}{\underset{j \quad 0}{}} \quad {}_j \quad 1.$$

The $_j$ are called the **steady-state probabilities** of the Markov chain. The term *steady-state* probability means that the probability of finding the process in a certain state, say j, after a large number of transitions tends to the value $_j$, independent of the probability distribution of the initial state. It is important to note that the steady-state probability does *not* imply that the process settles down into one state. On the contrary, the process continues to make transitions from state to state, and at any step n the transition probability from state i to state j is still p_{ij}.

The $_j$ can also be interpreted as *stationary probabilities* (not to be confused with stationary transition probabilities) in the following sense. If the *initial* probability of being in state j is given by $_j$ (that is, $P\{X_0 \quad j\} \quad {}_j$) for all j, then the probability of finding the process in state j at time n 1, 2, . . . is also given by $_j$ (that is, $P\{X_n \quad j\} \quad {}_j$).

Note that the steady-state equations consist of M 2 equations in M 1 unknowns. Because it has a unique solution, at least one equation must be redundant and can, therefore, be deleted. It cannot be the equation

$$\overset{M}{\underset{j \quad 0}{}} \quad {}_j \quad 1,$$

because $_j$ 0 for all j will satisfy the other M 1 equations. Furthermore, the solutions to the other M 1 steady-state equations have a unique solution up to a multiplicative constant, and it is the final equation that forces the solution to be a probability distribution.

Returning to the inventory example, we see that the steady-state equations can be expressed as

$$
\begin{array}{lllll}
_0 & {}_0 p_{00} & {}_1 p_{10} & {}_2 p_{20} & {}_3 p_{30}, \\
_1 & {}_0 p_{01} & {}_1 p_{11} & {}_2 p_{21} & {}_3 p_{31}, \\
_2 & {}_0 p_{02} & {}_1 p_{12} & {}_2 p_{22} & {}_3 p_{32}, \\
_3 & {}_0 p_{03} & {}_1 p_{13} & {}_2 p_{23} & {}_3 p_{33}, \\
1 & {}_0 & {}_1 & {}_2 & {}_3 .
\end{array}
$$

Substituting values for p_{ij} into these equations leads to the equations

$$
\begin{aligned}
{}_0 &\quad 0.080\,{}_0 \quad 0.632\,{}_1 \quad 0.264\,{}_2 \quad 0.080\,{}_3,\\
{}_1 &\quad 0.184\,{}_0 \quad 0.368\,{}_1 \quad 0.368\,{}_2 \quad 0.184\,{}_3,\\
{}_2 &\quad 0.368\,{}_0 \qquad\qquad\quad 0.368\,{}_2 \quad 0.368\,{}_3,\\
{}_3 &\quad 0.368\,{}_0 \qquad\qquad\qquad\qquad\quad 0.368\,{}_3,\\
1 &\qquad\quad {}_0 \qquad\quad {}_1 \qquad\quad {}_2 \qquad\quad {}_3.
\end{aligned}
$$

Solving the last four equations simultaneously provides the solution

$$
{}_0 \quad 0.286, \qquad {}_1 \quad 0.285, \qquad {}_2 \quad 0.263, \qquad {}_3 \quad 0.166,
$$

which is essentially the result that appears in matrix $\mathbf{P}^{(8)}$. Thus, after many weeks the probability of finding zero, one, two, and three cameras in stock tends to 0.286, 0.285, 0.263, and 0.166, respectively.

Your OR Courseware includes a routine for solving the steady-state equations to obtain the steady-state probabilities.

There are other important results concerning steady-state probabilities. In particular, if i and j are recurrent states belonging to different classes, then

$$
p_{ij}^{(n)} \quad 0, \qquad \text{for all } n.
$$

This result follows from the definition of a class.

Similarly, if j is a transient state, then

$$
\lim_{n\to} p_{ij}^{(n)} \quad 0, \qquad \text{for all } i.
$$

Thus, the probability of finding the process in a transient state after a large number of transitions tends to zero.

Expected Average Cost per Unit Time

The preceding subsection dealt with finite-state Markov chains whose states were ergodic (recurrent and aperiodic). If the requirement that the states be aperiodic is relaxed, then the limit

$$
\lim_{n\to} p_{ij}^{(n)}
$$

may not exist. To illustrate this point, consider the two-state transition matrix

$$
\mathbf{P} \quad
\begin{array}{c|cc}
\text{State} & 0 & 1\\
\hline
0 & 0 & 1\\
1 & 1 & 0
\end{array}.
$$

If the process starts in state 0 at time 0, it will be in state 0 at times 2, 4, 6, . . . and in state 1 at times 1, 3, 5, ... Thus, $p_{00}^{(n)} \quad 1$ if n is even and $p_{00}^{(n)} \quad 0$ if n is odd, so that

$$
\lim_{n\to} p_{00}^{(n)}
$$

does not exist. However, the following limit always exists for an irreducible (finite-state) Markov chain:

$$\lim_{n\to\infty}\left(\frac{1}{n}\sum_{k=1}^{n} p_{ij}^{(k)}\right) = \pi_j,$$

where the π_j satisfy the steady-state equations given in the preceding subsection.

This result is important in computing the *long-run average cost per unit time* associated with a Markov chain. Suppose that a cost (or other penalty function) $C(X_t)$ is incurred when the process is in state X_t at time t, for $t = 0, 1, 2, \ldots$. Note that $C(X_t)$ is a random variable that takes on any one of the values $C(0), C(1), \ldots, C(M)$ and that the function $C(\cdot)$ is independent of t. The expected average cost incurred over the first n periods is given by

$$E\left[\frac{1}{n}\sum_{t=1}^{n} C(X_t)\right].$$

By using the result that

$$\lim_{n\to\infty}\left(\frac{1}{n}\sum_{k=1}^{n} p_{ij}^{(k)}\right) = \pi_j,$$

it can be shown that the (long-run) *expected average cost per unit time* is given by

$$\lim_{n\to\infty} E\left[\frac{1}{n}\sum_{t=1}^{n} C(X_t)\right] = \sum_{j=0}^{M} \pi_j C(j).$$

To illustrate, consider the inventory example introduced in Sec. 16.1, where the solution for the π_j was obtained in the preceding subsection. Suppose the camera store finds that a storage charge is being allocated for each camera remaining on the shelf at the end of the week. The cost is charged as follows:

$$C(x_t) = \begin{cases} 0 & \text{if } x_t = 0 \\ 2 & \text{if } x_t = 1 \\ 8 & \text{if } x_t = 2 \\ 18 & \text{if } x_t = 3 \end{cases}$$

The long-run expected average storage cost per week can then be obtained from the preceding equation, i.e.,

$$\lim_{n\to\infty} E\left[\frac{1}{n}\sum_{t=1}^{n} C(X_t)\right] = 0.286(0) + 0.285(2) + 0.263(8) + 0.166(18) = 5.662.$$

Note that an alternative measure to the (long-run) expected average cost per unit time is the (long-run) *actual average cost per unit time.* It can be shown that this latter measure is given by

$$\lim_{n\to\infty}\left[\frac{1}{n}\sum_{t=1}^{n} C(X_t)\right] = \sum_{j=0}^{M} \pi_j C(j)$$

for essentially all paths of the process. Thus, either measure leads to the same result. These results can also be used to interpret the meaning of the π_j. To do so, let

$$C(X_t) = \begin{cases} 1 & \text{if } X_t = j \\ 0 & \text{if } X_t \neq j. \end{cases}$$

The (long-run) expected fraction of times the system is in state j is then given by

$$\lim_{n\to\infty} E\left(\frac{1}{n}\sum_{t=1}^{n} C(X_t)\right) = \lim_{n\to\infty} E(\text{fraction of times system is in state } j) = \pi_j.$$

Similarly, π_j can also be interpreted as the (long-run) actual fraction of times that the system is in state j.

Expected Average Cost per Unit Time for Complex Cost Functions

In the preceding subsection, the cost function was based solely on the state that the process is in at time t. In many important problems encountered in practice, the cost may also depend upon some other random variable.

For example, in the inventory example of Sec. 16.1, suppose that the costs to be considered are the ordering cost and the penalty cost for unsatisfied demand (storage costs are so small they will be ignored). It is reasonable to assume that the number of cameras ordered to arrive at the beginning of week t depends only upon the state of the process X_{t-1} (the number of cameras in stock) when the order is placed at the end of week $t-1$. However, the cost of unsatisfied demand in week t will also depend upon the demand D_t. Therefore, the total cost (ordering cost plus cost of unsatisfied demand) for week t is a function of X_{t-1} and D_t, that is, $C(X_{t-1}, D_t)$.

Under the assumptions of this example, it can be shown that the (long-run) *expected average cost per unit time* is given by

$$\lim_{n\to\infty} E\left[\frac{1}{n}\sum_{t=1}^{n} C(X_{t-1}, D_t)\right] = \sum_{j=0}^{M} k(j)\pi_j,$$

where

$$k(j) = E[C(j, D_t)],$$

and where this latter (conditional) expectation is taken with respect to the probability distribution of the random variable D_t, given the state j. Similarly, the (long-run) actual average cost per unit time is given by

$$\lim_{n\to\infty} \left[\frac{1}{n}\sum_{t=1}^{n} C(X_{t-1}, D_t)\right] = \sum_{j=0}^{M} k(j)\pi_j.$$

Now let us assign numerical values to the two components of $C(X_{t-1}, D_t)$ in this example, namely, the ordering cost and the penalty cost for unsatisfied demand. If $z > 0$ cameras are ordered, the cost incurred is $(10 + 25z)$ dollars. If no cameras are ordered, no ordering cost is incurred. For each unit of unsatisfied demand (lost sales), there is a penalty of \$50. Therefore, given the ordering policy described in Sec. 16.1, the cost in week t is given by

$$C(X_{t\quad 1}, D_t) \quad \begin{matrix} 10 \quad (25)(3) \quad 50\max\{D_t \quad 3, 0\} & \text{if } X_{t\quad 1} \quad 0 \\ 50\max\{D_t \quad X_{t\quad 1}, 0\} & \text{if } X_{t\quad 1} \quad 1, \end{matrix}$$

for t 1, 2, ... Hence,

$$C(0, D_t) \quad 85 \quad 50\max\{D_t \quad 3, 0\},$$

so that

$$\begin{aligned} k(0) \quad & E[C(0, D_t)] \quad 85 \quad 50E(\max\{D_t \quad 3, 0\}) \\ & 85 \quad 50[P_D(4) \quad 2P_D(5) \quad 3P_D(6) \qquad], \end{aligned}$$

where $P_D(i)$ is the probability that the demand equals i, as given by a Poisson distribution with a mean of 1, so that $P_D(i)$ becomes negligible for i larger than about 6. Since $P_D(4)$ 0.015, $P_D(5)$ 0.003, and $P_D(6)$ 0.001, we obtain $k(0)$ 86.2. Also using $P_D(2)$ 0.184 and $P_D(3)$ 0.061, similar calculations lead to the results

$$\begin{aligned} k(1) \quad E[C(1, D_t)] \quad & 50E(\max\{D_t \quad 1, 0\}) \\ & 50[P_D(2) \quad 2P_D(3) \quad 3P_D(4) \qquad] \\ & 18.4, \\ k(2) \quad E[C(2, D_t)] \quad & 50E(\max\{D_t \quad 2, 0\}) \\ & 50[P_D(3) \quad 2P_D(4) \quad 3P_D(5) \qquad] \\ & 5.2, \end{aligned}$$

and

$$\begin{aligned} k(3) \quad E[C(3, D_t)] \quad & 50E(\max\{D_t \quad 3, 0\}) \\ & 50[P_D(4) \quad 2P_D(5) \quad 3P_D(6) \qquad] \\ & 1.2. \end{aligned}$$

Thus, the (long-run) expected average cost per week is given by

$$\sum_{j\quad 0}^{3} k(j) \quad _j \quad 86.2(0.286) \quad 18.4(0.285) \quad 5.2(0.263) \quad 1.2(0.166) \quad \$31.46.$$

This is the cost associated with the particular ordering policy described in Sec. 16.1. The cost of other ordering policies can be evaluated in a similar way to identify the policy that minimizes the expected average cost per week.

The results of this subsection were presented only in terms of the inventory example. However, the (nonnumerical) results still hold for other problems as long as the following conditions are satisfied:

1. $\{X_t\}$ is an irreducible (finite-state) Markov chain.
2. Associated with this Markov chain is a sequence of random variables $\{D_t\}$ which are independent and identically distributed.
3. For a fixed m 0, 1, 2, . . . , a cost $C(X_t, D_{t\quad m})$ is incurred at time t, for t 0, 1, 2, ...
4. The sequence $X_0, X_1, X_2, \ldots, X_t$ must be independent of $D_{t\quad m}$.

By Dr.S.K.Rath, BPUT

In particular, if these conditions are satisfied, then

$$\lim \ E \ \ 1 \ {}^{n}_{t \ \ 1} \ \ C(X_t, D_{t \ \ m}) \qquad {}^{M}_{j \ \ 0} \ k(j) \ \ {}_{j},$$

where

$$k(j) \qquad E[C(j, D_{t \ \ m})],$$

and where this latter conditional expectation is taken with respect to the probability distribution of the random variable D_t, given the state j. Furthermore,

$$\lim \ \ 1 \ {}^{n}_{t \ \ 1} \ \ C(X_t, D_{t \ \ m}) \qquad {}^{M}_{j \ \ 0} \ k(j) \ \ {}_{j}$$

for essentially all paths of the process.

16.6 FIRST PASSAGE TIMES

Section 16.3 dealt with finding n-step transition probabilities from state i to state j. It is often desirable to also make probability statements about the number of transitions made by the process in going from state i to state j *for the first time.* This length of time is called the **first passage time** in going from state i to state j. When $j \quad i$, this first passage time is just the number of transitions until the process returns to the initial state i. In this case, the first passage time is called the **recurrence time** for state i.

To illustrate these definitions, reconsider the inventory example introduced in Sec. 16.1, where X_t is the number of cameras on hand at the end of week t, where we start with $X_0 \quad 3$. Suppose that it turns out that

$$X_0 \quad 3, \quad X_1 \quad 2, \quad X_2 \quad 1, \quad X_3 \quad 0, \quad X_4 \quad 3, \quad X_5 \quad 1.$$

In this case, the first passage time in going from state 3 to state 1 is 2 weeks, the first passage time in going from state 3 to state 0 is 3 weeks, and the recurrence time for state 3 is 4 weeks.

In general, the first passage times are random variables. The probability distributions associated with them depend upon the transition probabilities of the process. In particular, let $f_{ij}^{(n)}$ denote the probability that the first passage time from state i to j is equal to n. For $n \quad 1$, this first passage time is n if the first transition is from state i to some state k ($k \quad j$) and then the first passage time from state k to state j is $n \quad 1$. Therefore, these probabilities satisfy the following recursive relationships:

$$f_{ij}^{(1)} \qquad p_{ij}^{(1)} \qquad p_{ij},$$

$$f_{ij}{}^{(2)} \quad {}_{k \ \ j} \ p_{ik} f_{kj}^{(1)},$$

$$f_{ij}{}^{(n)} \quad {}_{k \ \ j} \ p_{ik} f_{kj}^{(n \quad 1)}.$$

Thus, the probability of a first passage time from state i to state j in n steps can be computed recursively from the one-step transition probabilities.

In the inventory example, the probability distribution of the first passage time in going from state 3 to state 0 is obtained from these recursive relationships as follows:

$$
\begin{aligned}
f_{30}^{(1)} &= p_{30} = 0.080,\\
f_{30}^{(2)} &= p_{31} f_{10}^{(1)} + p_{32} f_{20}^{(1)} + p_{33} f_{30}^{(1)}\\
&= 0.184(0.632) + 0.368(0.264) + 0.368(0.080) = 0.243,
\end{aligned}
$$

where the p_{3k} and $f_{k0}^{(1)} = p_{k0}$ are obtained from the (one-step) transition matrix given in Sec. 16.2.

For fixed i and j, the $f_{ij}^{(n)}$ are nonnegative numbers such that

$$\sum_{n=1}^{\infty} f_{ij}^{(n)} \le 1.$$

Unfortunately, this sum may be strictly less than 1, which implies that a process initially in state i may never reach state j. When the sum does equal 1, $f_{ij}^{(n)}$ (for $n = 1, 2, \ldots$) can be considered as a probability distribution for the random variable, the first passage time.

Although obtaining $f_{ij}^{(n)}$ for all n may be tedious, it is relatively simple to obtain the expected first passage time from state i to state j. Denote this expectation by μ_{ij}, which is defined by

$$\mu_{ij} = \begin{cases} \infty & \text{if } \sum_{n=1}^{\infty} f_{ij}^{(n)} < 1 \\ \sum_{n=1}^{\infty} n f_{ij}^{(n)} & \text{if } \sum_{n=1}^{\infty} f_{ij}^{(n)} = 1. \end{cases}$$

Whenever

$$\sum_{n=1}^{\infty} f_{ij}^{(n)} = 1,$$

μ_{ij} uniquely satisfies the equation

$$\mu_{ij} = 1 + \sum_{k \ne j} p_{ik} \mu_{kj}.$$

This equation recognizes that the first transition from state i can be to either state j or to some other state k. If it is to state j, the first passage time is 1. Given that the first transition is to some state k ($k \ne j$) instead, which occurs with probability p_{ik}, the conditional expected first passage time from state i to state j is $1 + \mu_{kj}$. Combining these facts, and summing over all the possibilities for the first transition, leads directly to this equation.

For the inventory example, these equations for the μ_{ij} can be used to compute the expected time until the cameras are out of stock, given that the process is started when three cameras are available. This expected time is just the expected first passage time μ_{30}. Since all the states are recurrent, the system of equations leads to the expressions

$$
\begin{aligned}
\mu_{30} &= 1 + p_{31}\mu_{10} + p_{32}\mu_{20} + p_{33}\mu_{30},\\
\mu_{20} &= 1 + p_{21}\mu_{10} + p_{22}\mu_{20} + p_{23}\mu_{30},\\
\mu_{10} &= 1 + p_{11}\mu_{10} + p_{12}\mu_{20} + p_{13}\mu_{30},
\end{aligned}
$$

or

$$
\begin{aligned}
&_{30} \quad 1 \quad 0.184 \;_{10} \quad 0.368 \;_{20} \quad 0.368 \;_{30},\\
&_{20} \quad 1 \quad 0.368 \;_{10} \quad 0.368 \;_{20},\\
&_{10} \quad 1 \quad 0.368 \;_{10}.
\end{aligned}
$$

The simultaneous solution to this system of equations is

$$
\begin{aligned}
&_{10} \quad 1.58 \text{ weeks},\\
&_{20} \quad 2.51 \text{ weeks},\\
&_{30} \quad 3.50 \text{ weeks},
\end{aligned}
$$

so that the expected time until the cameras are out of stock is 3.50 weeks. Thus, in making these calculations for $_{30}$, we also obtain $_{20}$ and $_{10}$.

For the case of $_{ij}$ where $j \quad i$, $_{ii}$ is the expected number of transitions until the process returns to the initial state i, and so is called the **expected recurrence time** for state i. After obtaining the steady-state probabilities ($_0$, $_1$, . . . , $_M$) as described in the preceding section, these expected recurrence times can be calculated immediately as

$$_{ii} \quad \frac{1}{_i}, \qquad \text{for } i \quad 0, 1, \ldots, M.$$

Thus, for the inventory example, where $_0$ 0.286, $_1$ 0.285, $_2$ 0.263, and $_3$ 0.166, the corresponding expected recurrence times are

$$
\begin{aligned}
&_{00} \quad \frac{1}{_0} \quad 3.50 \text{ weeks}, \qquad &_{22} \quad \frac{1}{_2} \quad 3.80 \text{ weeks},\\
&_{11} \quad \frac{1}{_1} \quad 3.51 \text{ weeks}, \qquad &_{33} \quad \frac{1}{_3} \quad 6.02 \text{ weeks}.
\end{aligned}
$$

16.7 ABSORBING STATES

It was pointed out in Sec. 16.4 that a state k is called an *absorbing state* if p_{kk} 1, so that once the chain visits k it remains there forever. If k is an absorbing state, and the process starts in state i, the probability of *ever* going to state k is called the **probability of absorption** into state k, given that the system started in state i. This probability is denoted by f_{ik}.

When there are two or more absorbing states in a Markov chain, and it is evident that the process will be absorbed into one of these states, it is desirable to find these probabilities of absorption. These probabilities can be obtained by solving a system of linear equations that considers all the possibilities for the first transition and then, given the first transition, considers the conditional probability of absorption into state k. In particular, if the state k is an absorbing state, then the set of absorption probabilities f_{ik} satisfies the system of equations

$$f_{ik} \quad {}_{j \quad 0}^{M} \; p_{ij} f_{jk}, \qquad \text{for } i \quad 0, 1, \ldots, M,$$

subject to the conditions

$$
\begin{aligned}
&f_{kk} \quad 1,\\
&f_{ik} \quad 0, \qquad \text{if state } i \text{ is recurrent and } i \quad k.
\end{aligned}
$$

Absorption probabilities are important in random walks. A **random walk** is a Markov chain with the property that if the system is in a state i, then in a single transition the system either remains at i or moves to one of the two states immediately adjacent to i. For example, a random walk often is used as a model for situations involving gambling.

To illustrate, consider a gambling example similar to that presented in Sec. 16.2. However, suppose now that two players (A and B), each having \$2, agree to keep playing the game and betting \$1 at a time until one player is broke. The probability of A winning a single bet is $\frac{1}{3}$, so B wins the bet with probability $\frac{2}{3}$. The number of dollars that player A has before each bet (0, 1, 2, 3, or 4) provides the states of a Markov chain with transition matrix

$$\mathbf{P} = \begin{array}{c} \text{State} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}.$$

Starting from state 2, the probability of absorption into state 0 (A losing all her money) can be obtained from the preceding system of equations as $f_{20}\ \frac{1}{5}$, and the probability of A winning \$4 (B going broke) is given by $f_{24}\ \frac{4}{5}$.

There are many other situations where absorbing states play an important role. Consider a department store that classifies the balance of a customer's bill as fully paid (state 0), 1 to 30 days in arrears (state 1), 31 to 60 days in arrears (state 2), or bad debt (state 3). The accounts are checked *monthly* to determine the state of each customer. In general, credit is not extended and customers are expected to pay their bills within 30 days. Occasionally, customers pay only portions of their bill. If this occurs when the balance is within 30 days in arrears (state 1), the store views the customer as remaining in state 1. If this occurs when the balance is between 31 and 60 days in arrears, the store views the customer as moving to state 1 (1 to 30 days in arrears). Customers that are more than 60 days in arrears are put into the bad-debt category (state 3), and then bills are sent to a collection agency. After examining data over the past several years on the month by month progression of individual customers from state to state, the store has developed the following transition matrix:[1]

[1]Customers who are fully paid (in state 0) and then subsequently fall into arrears on new purchases are viewed as "new" customers who start in state 1.

State \ State	0: Fully Paid	1: 1 to 30 Days in Arrears	2: 31 to 60 Days in Arrears	3: Bad Debt
0: fully paid	1	0	0	0
1: 1 to 30 days in arrears	0.7	0.2	0.1	0
2: 31 to 60 days in arrears	0.5	0.1	0.2	0.2
3: bad debt	0	0	0	1

Although each customer ends up in state 0 or 3, the store is interested in determining the probability that a customer will end up as a bad debt given that the account belongs to the 1 to 30 days in arrears state, and similarly, given that the account belongs to the 31 to 60 days in arrears state.

To obtain this information, the set of equations presented at the beginning of this section must be solved to obtain f_{13} and f_{23}. By substituting, the following two equations are obtained:

$$f_{13} \quad p_{10} f_{03} \quad p_{11} f_{13} \quad p_{12} f_{23} \quad p_{13} f_{33},$$
$$f_{23} \quad p_{20} f_{03} \quad p_{21} f_{13} \quad p_{22} f_{23} \quad p_{23} f_{33}.$$

Noting that $f_{03} \quad 0$ and $f_{33} \quad 1$, we now have two equations in two unknowns, namely,

$$(1 \quad p_{11}) f_{13} \quad p_{13} \quad p_{12} f_{23},$$
$$(1 \quad p_{22}) f_{23} \quad p_{23} \quad p_{21} f_{13}.$$

Substituting the values from the transition matrix leads to

$$0.8f_{13} \quad 0.1 f_{23},$$
$$0.8f_{23} \quad 0.2 \quad 0.1 f_{13},$$

and the solution is

$$f_{13} \quad 0.032,$$
$$f_{23} \quad 0.254.$$

Thus, approximately 3 percent of the customers whose accounts are 1 to 30 days in arrears end up as bad debts, whereas about 25 percent of the customers whose accounts are 31 to 60 days in arrears end up as bad debts.

16.8 CONTINUOUS TIME MARKOV CHAINS

In all the previous sections, we assumed that the time parameter t was discrete (that is, $t \quad 0, 1, 2, \ldots$). Such an assumption is suitable for many problems, but there are certain cases (such as for some queueing models considered in the next chapter) where a continuous time parameter (call it t) is required, because the evolution of the process is being observed *continuously* over time. The definition of a Markov chain given in Sec. 16.2 also extends to such continuous processes. This section focuses on describing these "continuous time Markov chains" and their properties.

Formulation

As before, we label the possible **states** of the system as $0, 1, \ldots, M$. Starting at time 0 and letting the time parameter t run continuously for $t \quad 0$, we let the random variable $X(t)$ be the state of the system at time t. Thus, $X(t)$ will take on one of its possible $(M \quad 1)$ values over some interval, $0 \quad t \quad t_1$, then will jump to another value over the next interval, $t_1 \quad t \quad t_2$, etc., where these transit points $(t_1, t_2, \ldots)$ are random points in time (*not* necessarily integer).

Now consider the three points in time (1) $t \quad r$ (where $r \quad 0$), (2) $t \quad s$ (where $s \quad r$), and (3) $t \quad s \quad t$ (where $t \quad 0$), interpreted as follows:

$t \quad r$ is a past time,
$t \quad s$ is the current time,
$t \quad s \quad t$ is t time units into the future.

Therefore, the state of the system now has been observed at times t s and t r. Label these states as

$$X(s) \quad i \qquad \text{and} \qquad X(r) \quad x(r).$$

Given this information, it now would be natural to seek the probability distribution of the state of the system at time t s t. In other words, what is

$$P\{X(s \quad t) \quad j \mid X(s) \quad i \text{ and } X(r) \quad x(r)\}, \qquad \text{for } j \quad 0, 1, \ldots, M?$$

Deriving this conditional probability often is very difficult. However, this task is considerably simplified if the stochastic process involved possesses the following key property.

A continuous time stochastic process $\{X(t); t \quad 0\}$ has the **Markovian property** if

$$P\{X(t \quad s) \quad j \mid X(s) \quad i \text{ and } X(r) \quad x(r)\} \quad P\{X(t \quad s) \quad j \mid X(s) \quad i\},$$

for all i, j $0, 1, \ldots, M$ and for all r 0, s r, and t 0.

Note that $P\{X(t \quad s) \quad j \mid X(s) \quad i\}$ is a **transition probability,** just like the transition probabilities for discrete time Markov chains considered in the preceding sections, where the only difference is that t now need not be an integer.

If the transition probabilities are independent of s, so that

$$P\{X(t \quad s) \quad j \mid X(s) \quad i\} \quad P\{X(t) \quad j \mid X(0) \quad i\}$$

for all s 0, they are called **stationary transition probabilities.**

To simplify notation, we shall denote these stationary transition probabilities by

$$p_{ij}(t) \quad P\{X(t) \quad j \mid X(0) \quad i\},$$

where $p_{ij}(t)$ is referred to as the **continuous time transition probability function.** We assume that

$$\lim_{t \to 0} p_{ij}(t) \quad \begin{matrix} 1 & \text{if } i \quad j \\ 0 & \text{if } i \quad j. \end{matrix}$$

Now we are ready to define the continuous time Markov chains to be considered in this section.

A continuous time stochastic process $\{X(t); t \quad 0\}$ is a **continuous time Markov chain** if it has the *Markovian property.*

We shall restrict our consideration to continuous time Markov chains with the following properties:

1. A finite number of states.
2. Stationary transition probabilities.

Some Key Random Variables

In the analysis of continuous time Markov chains, one key set of random variables is the following.

Each time the process enters state i, the amount of time it spends in that state before moving to a different state is a random variable T_i, where i $0, 1, \ldots, M$.

Suppose that the process enters state i at time $t \quad s$. Then, for any fixed amount of time $t \quad 0$, note that $T_i \quad t$ if and only if $X(t\ ) \quad i$ for all t over the interval $s \quad t \quad s \quad t$. Therefore, the Markovian property (with stationary transition probabilities) implies that

$$P\{T_i \quad t \quad s \mid T_i \quad s\} \quad P\{T_i \quad t\}.$$

This is a rather unusual property for a probability distribution to possess. It says that the probability distribution of the *remaining* time until the process transits out of a given state always is the same, regardless of how much time the process has already spent in that state. In effect, the random variable is memoryless; the process forgets its history. There is only one (continuous) probability distribution that possesses this property—the *exponential distribution*. The exponential distribution has a single parameter, call it q, where the mean is $1/q$ and the cumulative distribution function is

$$P\{T_i \quad t\} \quad 1 \quad e^{\quad qt}, \qquad \text{for } t \quad 0.$$

(We shall describe the properties of the exponential distribution in detail in Sec. 17.4.)

This result leads to an equivalent way of describing a continuous time Markov chain:

1. The random variable T_i has an exponential distribution with a mean of $1/q_i$.

2. When leaving state i, the process moves to a state j with probability p_{ij}, where the p_{ij} satisfy the conditions

$$p_{ij} \quad 0 \qquad \text{for all } i,$$

and

$$\sum_{j \quad 0}^{M} p_{ij} \quad 1 \qquad \text{for all } i.$$

3. The next state visited after state i is independent of the time spent in state i.

Just as the one-step transition probabilities played a major role in describing discrete time Markov chains, the analogous role for a continuous time Markov chain is played by the transition intensities.

The **transition intensities** are

$$q_i \quad \frac{d}{dt} p_{ii}(0) \quad \lim_{t \to 0} \frac{1 \quad p_{ii}(t)}{t}, \qquad \text{for } i \quad 0, 1, 2, \ldots, M,$$

and

$$q_{ij} \quad \frac{d}{dt} p_{ij}(0) \quad \lim_{t \to 0} \frac{p_{ij}(t)}{t} \quad q_i p_{ij}, \qquad \text{for all } j \quad i,$$

where $p_{ij}(t)$ is the *continuous time transition probability function* introduced at the beginning of the section and p_{ij} is the probability described in property 2 of the preceding paragraph. Furthermore, q_i as defined here turns out to still be the parameter of the exponential distribution for T_i as well (see property 1 of the preceding paragraph).

The intuitive interpretation of the q_i and q_{ij} is that they are *transition rates*. In particular, q_i is the *transition rate out of state* i in the sense that q_i is the expected number of times that the process leaves state i per unit of time spent in state i. (Thus, q_i is the reciprocal of the expected time that the process spends in state i per visit to state i; that is, $q_i = 1/E[T_i]$.) Similarly, q_{ij} is the *transition rate from state i to state j* in the sense that q_{ij} is the expected number of times that the process transits from state i to state j per unit of time spent in state i. Thus,

$$q_i = \sum_{j \neq i} q_{ij}.$$

Just as q_i is the parameter of the exponential distribution for T_i, each q_{ij} is the parameter of an exponential distribution for a related random variable described below.

> Each time the process enters state i, the amount of time it will spend in state i before a transition to state j occurs (if a transition to some other state does not occur first) is a random variable T_{ij}, where $i, j = 0, 1, \ldots, M$ and $j \neq i$. The T_{ij} are independent random variables, where each T_{ij} has an *exponential distribution* with parameter q_{ij}, so $E[T_{ij}] = 1/q_{ij}$. The time spent in state i until a transition occurs (T_i) is the *minimum* (over $j \neq i$) of the T_{ij}. When the transition occurs, the probability that it is to state j is $p_{ij} = q_{ij}/q_i$.

Steady-State Probabilities

Just as the transition probabilities for a discrete time Markov chain satisfy the Chapman-Kolmogorov equations, the continuous time transition probability function also satisfies these equations. Therefore, for any states i and j and nonnegative numbers t and s $(0 \leq s \leq t)$,

$$p_{ij}(t) = \sum_{k=1}^{M} p_{ik}(s)p_{kj}(t - s).$$

A pair of states i and j are said to *communicate* if there are times t_1 and t_2 such that $p_{ij}(t_1) > 0$ and $p_{ji}(t_2) > 0$. All states that communicate are said to form a *class*. If all states form a single class, i.e., if the Markov chain is *irreducible* (hereafter assumed), then

$$p_{ij}(t) > 0, \qquad \text{for all } t > 0 \text{ and all states } i \text{ and } j.$$

Furthermore,

$$\lim_{t \to \infty} p_{ij}(t) = \pi_j$$

always exists and is independent of the initial state of the Markov chain, for $j = 0, 1, \ldots, M$. These limiting probabilities are commonly referred to as the **steady-state probabilities** (or *stationary probabilities*) of the Markov chain.

The π_j satisfy the equations

$$\pi_j = \sum_{i=0}^{M} \pi_i p_{ij}(t), \qquad \text{for } j = 0, 1, \ldots, M \text{ and every } t \geq 0.$$

However, the following **steady-state equations** provide a more useful system of equations for solving for the steady-state probabilities:

$$\pi_j q_j = \sum_{i \neq j} \pi_i q_{ij}, \qquad \text{for } j = 0, 1, \ldots, M.$$

and

$$\sum_{j=0}^{M} \pi_j = 1.$$

The steady-state equation for state j has an intuitive interpretation. The left-hand side ($\pi_j q_j$) is the *rate* at which the process *leaves* state j, since π_j is the (steady-state) probability that the process is in state j and q_j is the transition rate out of state j given that the process is in state j. Similarly, each term on the right-hand side ($\pi_i q_{ij}$) is the *rate* at which the process *enters* state j from state i, since q_{ij} is the transition rate from state i to state j given that the process is in state i. By summing over all $i \neq j$, the entire right-hand side then gives the rate at which the process enters state j from any other state. The overall equation thereby states that the rate at which the process leaves state j must equal the rate at which the process enters state j. Thus, this equation is analogous to the conservation of flow equations encountered in many engineering and science courses.

Because each of the first $M + 1$ *steady-state equations* requires that two rates be *in balance* (equal), these equations sometimes are called the **balance equations.**

Example. A certain shop has two identical machines that are operated continuously except when they are broken down. Because they break down fairly frequently, the top-priority assignment for a full-time maintenance person is to repair them whenever needed.

The time required to repair a machine has an exponential distribution with a mean of $\frac{1}{2}$ day. Once the repair of a machine is completed, the time until the next breakdown of that machine has an exponential distribution with a mean of 1 day. These distributions are independent.

Define the random variable $X(t)$ as

$$X(t) = \text{number of machines broken down at time } t,$$

so the possible values of $X(t)$ are 0, 1, 2. Therefore, by letting the time parameter t run continuously from time 0, the continuous time stochastic process $\{X(t); t \geq 0\}$ gives the evolution of the number of machines broken down.

Because both the repair time and the time until a breakdown have exponential distributions, $\{X(t); t \geq 0\}$ is a *continuous time Markov chain*[1] with states 0, 1, 2. Consequently, we can use the steady-state equations given in the preceding subsection to find the steady-state probability distribution of the number of machines broken down. To do this, we need to determine all the *transition rates,* i.e., the q_i and q_{ij} for $i, j = 0, 1, 2$.

The state (number of machines broken down) increases by 1 when a breakdown occurs and decreases by 1 when a repair occurs. Since both breakdowns and repairs occur one at a time, $q_{02} = 0$ and $q_{20} = 0$. The expected repair time is $\frac{1}{2}$ day, so the rate at which repairs are completed (when any machines are broken down) is 2 per day, which implies that $q_{21} = 2$ and $q_{10} = 2$. Similarly, the expected time until a particular operational machine breaks down is 1 day, so the rate at which it breaks down (when operational) is 1

[1]Proving this fact requires the use of two properties of the exponential distribution discussed in Sec. 17.4 (*lack of memory* and *the minimum of exponentials is exponential*), since these properties imply that the T_{ij} random variables introduced earlier do indeed have exponential distributions.

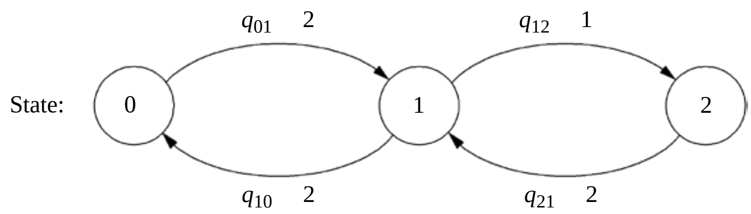


FIGURE 16.2
Rate diagram for the example of a continuous time Markov chain.

per day, which implies that q_{12} 1. During times when both machines are operational, breakdowns occur at the rate of 1 1 2 per day, so q_{01} 2.

These transition rates are summarized in the rate diagram shown in Fig. 16.2. These rates now can be used to calculate the *total transition rate* out of each state.

q_0 q_{01} 2.
q_1 q_{10} q_{12} 3.
q_2 q_{21} 2.

Plugging all the rates into the steady-state equations given in the preceding subsection then yields

Balance equation for state 0: $2_{\ 0}$ $2_{\ 1}$
Balance equation for state 1: $3_{\ 1}$ $2_{\ 0}$ $2_{\ 2}$
Balance equation for state 2: $2_{\ 2}$ $_{1}$
Probabilities sum to 1: $_{0}$ $_{1}$ $_{2}$ 1

Any one of the balance equations (say, the second) can be deleted as redundant, and the simultaneous solution of the remaining equations gives the steady-state distribution as

$$(\ _0,\ _1,\ _2)\ \ \frac{2}{5}, \frac{2}{5}, \frac{1}{5}\ .$$

Thus, in the long run, both machines will be broken down simultaneously 20 percent of the time, and one machine will be broken down another 40 percent of the time.

The next chapter (on queueing theory) features many more examples of continuous time Markov chains. In fact, most of the basic models of queueing theory fall into this category. The current example actually fits one of these models (the finite calling population variation of the *M*/*M*/*s* model included in Sec. 17.6).

SELECTED REFERENCES

1. Heyman, D., and M. Sobel: *Stochastic Models in Operations Research*, vol. 1, McGraw-Hill, New York, 1982.
2. Kao, E. P. C.: *An Introduction to Stochastic Processes*, Duxbury Press, Belmont, CA, 1997.
3. Resnick, S. I.: *Adventures in Stochastic Processes*, Birkhäuser, Boston, 1992.
4. Ross, S.: *Stochastic Processes*, 2d ed., Wiley, New York, 1995.

5. Stewart, W. J. (ed.): *Numerical Solution of Markov Chains,* Marcel Dekker, New York, 1991.
6. Taylor, H., and S. Karlin: *An Introduction to Stochastic Modeling,* revised edition, Academic Press, San Diego, 1993.
7. Tijms, H. C.: *Stochastic Models: An Algorithmic Approach,* Wiley, New York, 1994.

LEARNING AIDS FOR THIS CHAPTER IN YOUR OR COURSEWARE

Automatic Routines in OR Courseware:

Enter Transition Matrix
Chapman-Kolmogorov Equations
Steady-State Probabilities

See Appendix 1 for documentation of the software.

PROBLEMS

The symbol to the left of some of the problems (or their parts) has the following meaning.

C: Use the computer with the corresponding automatic routines listed above (or other equivalent routines) to solve the problem.

An asterisk on the problem number indicates that at least a partial answer is given in the back of the book.

16.2-1. Assume that the probability of rain tomorrow is 0.5 if it is raining today, and assume that the probability of its being clear (no rain) tomorrow is 0.9 if it is clear today. Also assume that these probabilities do not change if information is also provided about the weather before today.
(a) Explain why the stated assumptions imply that the *Markovian property* holds for the evolution of the weather.
(b) Formulate the evolution of the weather as a Markov chain by defining its states and giving its (one-step) transition matrix.

16.2-2. Consider the second version of the stock market model presented as an example in Sec. 16.2. Whether the stock goes up tomorrow depends upon whether it increased today *and* yesterday. If the stock increased today and yesterday, it will increase tomorrow with probability $_1$. If the stock increased today and decreased yesterday, it will increase tomorrow with probability $_2$. If the stock decreased today and increased yesterday, it will increase tomorrow with probability $_3$. Finally, if the stock decreased today and yesterday, it will increase tomorrow with probability $_4$.
(a) Construct the (one-step) transition matrix of the Markov chain.
(b) Explain why the states used for this Markov chain cause the mathematical definition of the Markovian property to hold even though what happens in the future (tomorrow) depends upon what happened in the past (yesterday) as well as the present (today).

16.2-3. Reconsider Prob. 16.2-2. Suppose now that whether or not the stock goes up tomorrow depends upon whether it increased today, yesterday, *and* the day before yesterday. Can this problem be formulated as a Markov chain? If so, what are the possible states? Explain why these states give the process the *Markovian property* whereas the states in Prob. 16.2-2 do not.

16.3-1. Reconsider Prob. 16.2-1.
C **(a)** Use the routine *Chapman-Kolmogorov Equations* in your OR Courseware to find the n-step transition matrix $\mathbf{P}^{(n)}$ for n 2, 5, 10, 20.
(b) The probability that it will rain today is 0.5. Use the results from part (*a*) to determine the probability that it will rain n days from now, for n 2, 5, 10, 20.
C **(c)** Use the routine *Steady-State Probabilities* in your OR Courseware to determine the steady-state probabilities of the state of the weather. Describe how the probabilities in the n-step transition matrices obtained in part (*a*) compare to these steady-state probabilities as n grows large.

16.3-2. Suppose that a communications network transmits binary digits, 0 or 1, where each digit is transmitted 10 times in succession. During each transmission, the probability is 0.99 that the digit entered will be transmitted accurately. In other words, the probability is 0.01 that the digit being transmitted will be recorded with the opposite value at the end of the transmission. For each transmission after the first one, the digit entered for transmission is the one that was recorded at the end of the preceding transmission. If X_0 denotes the binary digit entering the system, X_1 the binary digit recorded after the first transmission, X_2 the binary digit recorded after the second transmission, . . . , then $\{X_n\}$ is a Markov chain.

(a) Construct the (one-step) transition matrix.

C **(b)** Use your OR Courseware to find the 10-step transition matrix $\mathbf{P}^{(10)}$. Use this result to identify the probability that a digit entering the network will be recorded accurately after the last transmission.

C **(c)** Suppose that the network is redesigned to improve the probability that a single transmission will be accurate from 0.99 to 0.999. Repeat part (*b*) to find the new probability that a digit entering the network will be recorded accurately after the last transmission.

16.3-3.* A particle moves on a circle through points that have been marked 0, 1, 2, 3, 4 (in a clockwise order). The particle starts at point 0. At each step it has probability 0.5 of moving one point clockwise (0 follows 4) and 0.5 of moving one point counterclockwise. Let X_n (n 0) denote its location on the circle after step n. $\{X_n\}$ is a Markov chain.

(a) Construct the (one-step) transition matrix.

C **(b)** Use your OR Courseware to determine the n-step transition matrix $\mathbf{P}^{(n)}$ for n 5, 10, 20, 40, 80.

C **(c)** Use your OR Courseware to determine the steady-state probabilities of the state of the Markov chain. Describe how the probabilities in the n-step transition matrices obtained in part (*b*) compare to these steady-state probabilities as n grows large.

16.4-1.* Given the following (one-step) transition matrices of a Markov chain, determine the classes of the Markov chain and whether they are recurrent.

(a) P

State	0	1	2	3
0	0	0	$\frac{1}{3}$	$\frac{2}{3}$
1	1	0	0	0
2	0	1	0	0

(b) P 3 State 1 0 0

16.4-2. Given each of the following (one-step) transition matrices of a Markov chain, determine the classes of the Markov chain and whether they are recurrent.

(a) P

State	0	1	2	3
0	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
1	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$
2	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{1}{3}$
3	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0

(b) P

State	0	1	2
0	0	0	1
1	$\frac{1}{2}$	$\frac{1}{2}$	0
2	0	1	0

16.4-3. Given the following (one-step) transition matrix of a Markov chain, determine the classes of the Markov chain and whether they are recurrent.

P

State	0	1	2	3	4
0	$\frac{1}{4}$	$\frac{3}{4}$	0	0	0
1	$\frac{3}{4}$	$\frac{1}{4}$	0	0	0
2	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0	0
3	0	0	0	$\frac{3}{4}$	$\frac{1}{4}$
4	0	0	0	$\frac{1}{4}$	$\frac{3}{4}$

16.4-4. Determine the period of each of the states in the Markov chain that has the following (one-step) transition matrix.

P

State	0	1	2	3	4	5
0	0	0	0	$\frac{2}{3}$	0	$\frac{1}{3}$
1	0	0	1	0	0	0
2	1	0	0	0	0	0
3	0	$\frac{1}{4}$	0	0	$\frac{3}{4}$	0
4	0	0	1	0	0	0
5	0	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0

16.4-5. Consider the Markov chain that has the following (one-step) transition matrix.

State	0	1	2	3	4
0	0	$\frac{4}{5}$	0	$\frac{1}{5}$	0

1 $\frac{1}{?}$ 0 $\frac{1}{?}$ 0 0 0 1 2 3

[1 0 0 0] $\frac{1}{2}$ $\frac{1}{2}$

By Dr.S.K.Rath, BPUT

1 4 2 0

P 2 1 1 2

3 2 10 5

0

4

0

1 1 1

3 3 3

2 0 1 2 1 2 0

3 1 0 0 1

2 2

4

0 0

0 1 0

0

0

(a) Determine the classes of this Markov chain and, for each class, determine whether it is recurrent or transient.

(b) For each of the classes identified in part (*b*), determine the period of the states in that class.

16.5-1. Reconsider Prob. 16.2-1. Suppose now that the given probabilities, 0.5 and 0.9, are replaced by arbitrary values, and , respectively. Solve for the *steady-state probabilities* of the state of the weather in terms of and .

16.5-2. A transition matrix **P** is said to be doubly stochastic if the sum over each column equals 1; that is,

$$\sum_{i \; 0}^{M} p_{ij} \quad 1, \qquad \text{for all } j.$$

If such a chain is irreducible, aperiodic, and consists of M 1 states, show that

$$_{j} \quad \frac{1}{M \quad 1}, \qquad \text{for } j \quad 0, 1, \ldots, M.$$

16.5-3. Reconsider Prob. 16.3-3. Use the results given in Prob. 16.5-2 to find the steady-state probabilities for this Markov chain. Then find what happens to these steady-state probabilities if, at each step, the probability of moving one point clockwise changes to 0.9 and the probability of moving one point counterclockwise changes to 0.1.

C **16.5-4.** The leading brewery on the West Coast (labeled *A*) has hired an OR analyst to analyze its market position. It is particularly concerned about its major competitor (labeled *B*). The analyst believes that brand switching can be modeled as a Markov chain using three states, with states *A* and *B* representing customers drinking beer produced from the aforementioned breweries and state *C* representing all other brands. Data are taken monthly, and the analyst has constructed the following (one-step) transition matrix from past data.

	A	*B*	*C*
A	0.7	0.2	0.1
B	0.2	0.75	0.05
C	0.1	0.1	0.8

What are the steady-state market shares for the two major breweries?

16.5-5. Consider the following blood inventory problem facing a hospital. There is need for a rare blood type, namely, type AB, Rh negative blood. The demand *D* (in pints) over any 3-day period is given by

$P\{D \quad 0\} \quad 0.4, \qquad P\{D \quad 1\} \quad 0.3,$
$P\{D \quad 2\} \quad 0.2, \qquad \text{and} \qquad P\{D \quad 3\} \quad 0.1.$

Note that the expected demand is 1 pint, since $E(D)$ 0.3(1) 0.2(2) 0.1(3) 1. Suppose that there are 3 days between deliveries. The hospital proposes a policy of receiving 1 pint at each delivery and using the oldest blood first. If more blood is required than is on hand, an expensive emergency delivery is made. Blood is discarded if it is still on the shelf after 21 days. Denote the state of the system as the number of pints on hand just after a delivery. Thus, because of the discarding policy, the largest possible state is 7.

(a) Construct the (one-step) transition matrix for this Markov chain.

C **(b)** Find the steady-state probabilities of the state of the Markov chain.

(c) Use the results from part (*b*) to find the steady-state probability that a pint of blood will need to be discarded during a 3-day period. (*Hint*: Because the oldest blood is used first, a pint reaches 21 days only if the state was 7 and then *D* 0.)

(d) Use the results from part (*b*) to find the steady-state probability that an emergency delivery will be needed during the 3-day period between regular deliveries.

16.5-6. A soap company specializes in a luxury type of bath soap. The sales of this soap fluctuate between two levels—"Low" and "High"—depending upon two factors: (1) whether they advertise, and (2) the advertising and marketing of new products being done by competitors. The second factor is out of the company's control, but it is trying to determine what its own advertising policy should be. For example, the marketing manager's proposal is to advertise when sales are low but not to advertise when sales are high. Advertising in any quarter of a year has its primary impact on sales in the *following* quarter. Therefore, at the beginning of each quarter, the needed information is available to forecast accurately whether sales will be low or high that quarter and to decide whether to advertise that quarter.

The cost of advertising is $1 million for each quarter of a year in which it is done. When advertising is done during a quarter, the probability of having high sales the next quarter is $\frac{1}{2}$ or $\frac{3}{4}$, depending upon whether the current quarter's sales are low or high. These probabilities go down to $\frac{1}{4}$ or $\frac{1}{2}$ when advertising is not done during the current quarter. The company's quarterly profits (excluding advertising costs) are $4 million when sales are high but only $2 million when sales are low. (Hereafter, use units of millions of dollars.)

(a) Construct the (one-step) transition matrix for each of the following advertising strategies: (i) never advertise, (ii) always advertise, (iii) follow the marketing manager's proposal.

(b) Determine the steady-state probabilities manually for each of the three cases in part (*a*).

(c) Find the long-run expected average profit (including a deduction for advertising costs) per quarter for each of the three advertising strategies in part (*a*). Which of these strategies is best according to this measure of performance?

C **16.5-7.** In the last subsection of Sec. 16.5, the (long-run) expected average cost per week (based on just ordering costs and unsatisfied demand costs) is calculated for the inventory example of Sec. 16.1. Suppose now that the ordering policy is changed to the following. Whenever the number of cameras on hand at the end of the week is 0 or 1, an order is placed that will bring this number up to 3. Otherwise, no order is placed.

Recalculate the (long-run) expected average cost per week under this new inventory policy.

16.5-8.* Consider the inventory example introduced in Sec. 16.1, but with the following change in the ordering policy. If the number of cameras on hand at the end of each week is 0 or 1, two additional cameras will be ordered. Otherwise, no ordering will take place. Assume that the storage costs are the same as given in the second subsection of Sec. 16.5.

C **(a)** Find the steady-state probabilities of the state of this Markov chain.

(b) Find the long-run expected average storage cost per week.

16.5-9. Consider the following inventory policy for a certain product. If the demand during a period exceeds the number of items available, this unsatisfied demand is backlogged; i.e., it is filled when the next order is received. Let Z_n (n 0, 1, . . .) denote the amount of inventory on hand minus the number of units backlogged before ordering at the end of period n (Z_0 0). If Z_n is zero or positive, no orders are backlogged. If Z_n is negative, then Z_n represents the number of backlogged units and no inventory is on hand. At the end of period n, if Z_n 1, an order is placed for $2m$ units, where m is the smallest integer such that Z_n $2m$ 1. Orders are filled immediately.

Let D_1, D_2, . . . , be the demand for a product in periods 1, 2, . . . , respectively. Assume that the D_n are independent and identically distributed random variables taking on the values, 0, 1, 2, 3, 4, each with probability $\frac{1}{5}$. Let X_n denote the amount of stock on hand *after* ordering at the end of period n (where X_0 2), so that

$$X_n \quad \begin{matrix} X_{n\ 1} \quad D_n \quad 2m & \text{if } X_{n\ 1} \quad D_n \quad 1 \\ X_{n\ 1} \quad D_n & \text{if } X_{n\ 1} \quad D_n \quad 1 \end{matrix} \qquad (n \quad 1, 2, \ldots),$$

when $\{X_n\}$ (n 0, 1, . . .) is a Markov chain. It has only two states, 1 and 2, because the only time that ordering will take place is when Z_n 0,   1,   2, or   3, in which case 2, 2, 4, and 4 units are ordered, respectively, leaving X_n 2, 1, 2, 1, respectively.

(a) Construct the (one-step) transition matrix.

(b) Use the steady-state equations to solve manually for the steady-state probabilities.

(c) Now use the result given in Prob. 16.5-2 to find the steady-state probabilities.

(d) Suppose that the ordering cost is given by (2 $2m$) if an order is placed and zero otherwise. The holding cost per period is Z_n if Z_n 0 and zero otherwise. The shortage cost per period is $4Z_n$ if Z_n 0 and zero otherwise. Find the (long-run) expected average cost per unit time.

16.5-10. An important unit consists of two components placed in parallel. The unit performs satisfactorily if one of the two components is operating. Therefore, only one component is operated at a time, but both components are kept operational (capable of being operated) as often as possible by repairing them as needed. An operating component breaks down in a given period with probability 0.2. When this occurs, the parallel component takes over, if it is operational, at the beginning of the next period. Only one component can be repaired at a time. The repair of a component starts at the beginning of the first available period and is completed at the end of the next period. Let X_t be a vector consisting of two elements U and V, where U represents the number of components that are operational at the end of period t and V represents the number of periods of repair that have been completed on components that are not yet operational. Thus, V 0 if U 2 or if U 1 and the repair of the nonoperational component is just getting under way. Because a repair takes two periods, V 1 if U 0 (since then one nonoperational component is waiting to begin repair while the other one is entering its second period of repair) or if U 1 and the nonoperational component is entering its second period of repair. Therefore, the state space consists of the four states (2, 0), (1, 0), (0, 1), and (1, 1). Denote these four states by 0, 1, 2, 3, respectively. $\{X_t\}$ (t 0, 1, . . .) is a Markov chain (assume that X_0 0) with the (one-step) transition matrix

P

State	0	1	2	3
0	0.8	0.2	0	0
1	0	0	0.2	0.8
2	0	1	0	0
3	0.8	0.2	0	0

C **(a)** What is the probability that the unit will be inoperable (because both components are down) after n periods, for n 2, 5, 10, 20?

C **(b)** What are the steady-state probabilities of the state of this Markov chain?

(c) If it costs $30,000 per period when the unit is inoperable (both components down) and zero otherwise, what is the (long-run) expected average cost per period?

16.6-1. A computer is inspected at the end of every hour. It is found to be either working (up) or failed (down). If the computer is found to be up, the probability of its remaining up for the next hour is 0.90. If it is down, the computer is repaired, which may require more than 1 hour. Whenever the computer is down (regardless of how long it has been down), the probability of its still being down 1 hour later is 0.35.

(a) Construct the (one-step) transition matrix for this Markov chain.

(b) Use the approach described in Sec. 16.6 to find the $_{ij}$ (the expected first passage time from state i to state j) for all i and j.

16.6-2. A manufacturer has a machine that, when operational at the beginning of a day, has a probability of 0.1 of breaking down sometime during the day. When this happens, the repair is done the next day and completed at the end of that day.

(a) Formulate the evolution of the status of the machine as a Markov chain by identifying three possible states at the end of each day, and then constructing the (one-step) transition matrix.

(b) Use the approach described in Sec. 16.6 to find the $_{ij}$ (the expected first passage time from state *i* to state *j*) for all *i* and *j*. Use these results to identify the expected number of full days that the machine will remain operational before the next breakdown after a repair is completed.

(c) Now suppose that the machine already has gone 20 full days without a breakdown since the last repair was completed. How does the expected number of full days *hereafter* that the machine will remain operational before the next breakdown compare with the corresponding result from part (*b*) when the repair had just been completed? Explain.

16.6-3. Reconsider Prob. 16.6-2. Now suppose that the manufacturer keeps a spare machine that only is used when the primary machine is being repaired. During a repair day, the spare machine has a probability of 0.1 of breaking down, in which case it is repaired the next day. Denote the state of the system by (*x*, *y*), where *x* and *y*, respectively, take on the values 1 or 0 depending upon whether the primary machine (*x*) and the spare machine (*y*) are operational (value of 1) or not operational (value of 0) at the end of the day. [*Hint*: Note that (0, 0) is not a possible state.]

chain.

(b) Find the *expected recurrence time* for the state (1, 0).

16.6-4. Consider the inventory example presented in Sec. 16.1 except that demand now has the following probability distribution:

$$P\{D \quad 0\} \quad \frac{1}{4}, \qquad P\{D \quad 2\} \quad \frac{1}{4},$$

$$P\{D \quad 1\} \quad \frac{1}{2}, \qquad P\{D \quad 3) \quad 0.$$

The ordering policy now is changed to ordering just 2 cameras at the end of the week if none are in stock. As before, no order is placed if there are any cameras in stock. Assume that there is one camera in stock at the time (the end of a week) the policy is instituted.

(a) Construct the (one-step) transition matrix.

C **(b)** Find the probability distribution of the state of this Markov chain *n* weeks after the new inventory policy is instituted, for *n* 2, 5, 10.

(c) Find the $_{ij}$ (the expected first passage time from state *i* to state *j*) for all *i* and *j*.

C **(d)** Find the steady-state probabilities of the state of this Markov chain.

(e) Assuming that the store pays a storage cost for each camera remaining on the shelf at the end of the week according to the function *C*(0) 0, *C*(1) \$2, and *C*(2) \$8, find the long-run expected average storage cost per week.

16.6-5. A production process contains a machine that deteriorates rapidly in both quality and output under heavy usage, so that it is inspected at the end of each day. Immediately after inspection, the condition of the machine is noted and classified into one of four possible states:

State	Condition
0	Good as new
1	Operable—minimum deterioration
2	Operable—major deterioration
3	Inoperable and replaced by a good-as-new machine

The process can be modeled as a Markov chain with its (one-step) transition matrix **P** given by

State	0	1	2	3
0	0	$\frac{7}{8}$	$\frac{1}{16}$	$\frac{1}{16}$
1	0	3	1	1
2	0	0	$\frac{1}{2}$	$\frac{1}{2}$
3	1	0	0	0

C **(a)** Find the steady-state probabilities.

(b) If the costs of being in states 0, 1, 2, 3, are 0, \$1,000, \$3,000, and \$6,000, respectively, what is the long-run expected average cost per day?

(c) Find the *expected recurrence time* for state 0 (i.e., the expected length of time a machine can be used before it must be replaced).

16.7-1. Consider the following gambler's ruin problem. A gambler bets \$1 on each play of a game. Each time, he has a probability *p* of winning and probability *q* 1 *p* of losing the dollar bet. He will continue to play until he goes broke or nets a fortune of *T* dollars. Let X_n denote the number of dollars possessed by the gambler after the *n*th play of the game. Then

$$X_{n \quad 1} \quad \begin{cases} X \quad 1 & \text{with probability } p \\ X_n \quad 1 & \end{cases} \quad \begin{matrix} \text{for } 0 \quad X_n \quad T, \\ \text{with probability } q \\ 1 \quad p \end{matrix}$$

$$X_{n \quad 1} \quad X_n,$$

By Dr.S.K.Rath, BPUT

for X_n 0 or T.

$\{X_n\}$ is a Markov chain. The gambler starts with X_0 dollars, where X_0 is a positive integer less than T.

(a) Construct the (one-step) transition matrix of the Markov chain.
(b) Find the classes of the Markov chain.
(c) Let T 3 and p 0.3. Using the notation of Sec. 16.7, find f_{10}, f_{1T}, f_{20}, f_{2T}.
(d) Let T 3 and p 0.7. Find f_{10}, f_{1T}, f_{20}, f_{2T}.

16.7-2. A video cassette recorder manufacturer is so certain of its quality control that it is offering a complete replacement warranty if a recorder fails within 2 years. Based upon compiled data, the company has noted that only 1 percent of its recorders fail during the first year, whereas 5 percent of the recorders that survive the first year will fail during the second year. The warranty does not cover replacement recorders.

(a) Formulate the evolution of the status of a recorder as a Markov chain whose states include two absorption states that involve needing to honor the warranty or having the recorder survive the warranty period. Then construct the (one-step) transition matrix.
(b) Use the approach described in Sec. 16.7 to find the probability that the manufacturer will have to honor the warranty.

16.8-1. Reconsider the example presented at the end of Sec. 16.8. Suppose now that a third machine, identical to the first two, has been added to the shop. The one maintenance person still must maintain all the machines.

(a) Develop the *rate diagram* for this Markov chain.
(b) Construct the *steady-state equations.*
(c) Solve these equations for the *steady-state probabilities.*

16.8-2. The state of a particular continuous time Markov chain is defined as the number of jobs currently at a certain work center, where a maximum of three jobs are allowed. Jobs arrive individually. Whenever fewer than three jobs are present, the time until the next arrival has an exponential distribution with a mean of $\frac{1}{2}$ day. Jobs are processed at the work center one at a time and then leave immediately. Processing times have an exponential distribution with a mean of $\frac{1}{4}$ day.

(a) Construct the *rate diagram* for this Markov chain.
(b) Write the *steady-state equations.*
(c) Solve these equations for the *steady-state probabilities.*

17

Queueing Theory

Queues (waiting lines) are a part of everyday life. We all wait in queues to buy a movie ticket, make a bank deposit, pay for groceries, mail a package, obtain food in a cafeteria, start a ride in an amusement park, etc. We have become accustomed to considerable amounts of waiting, but still get annoyed by unusually long waits.

However, having to wait is not just a petty personal annoyance. The amount of time that a nation's populace wastes by waiting in queues is a major factor in both the quality of life there and the efficiency of the nation's economy. For example, before its dissolution, the U.S.S.R. was notorious for the tremendously long queues that its citizens frequently had to endure just to purchase basic necessities. Even in the United States today, it has been estimated that Americans spend 37,000,000,000 hours per year waiting in queues. If this time could be spent productively instead, it would amount to nearly 20 million person-years of useful work each year!

Even this staggering figure does not tell the whole story of the impact of causing excessive waiting. Great inefficiencies also occur because of other kinds of waiting than people standing in line. For example, making *machines* wait to be repaired may result in lost production. *Vehicles* (including ships and trucks) that need to wait to be unloaded may delay subsequent shipments. *Airplanes* waiting to take off or land may disrupt later travel schedules. Delays in *telecommunication* transmissions due to saturated lines may cause data glitches. Causing *manufacturing jobs* to wait to be performed may disrupt subsequent production. Delaying *service jobs* beyond their due dates may result in lost future business.

Queueing theory is the study of waiting in all these various guises. It uses *queueing models* to represent the various types of *queueing systems* (systems that involve queues of some kind) that arise in practice. Formulas for each model indicate how the corresponding queueing system should perform, including the average amount of waiting that will occur, under a variety of circumstances.

Therefore, these queueing models are very helpful for determining how to operate a queueing system in the most effective way. Providing too much service capacity to operate the system involves excessive costs. But not providing enough service capacity results in excessive waiting and all its unfortunate consequences. The models enable finding an appropriate balance between the cost of service and the amount of waiting.

After some general discussion, this chapter presents most of the more elementary queueing models and their basic results. Chapter 18 discusses how the information provided by queueing theory can be used to design queueing systems that minimize the total cost of service and waiting.

17.1 PROTOTYPE EXAMPLE

The emergency room of COUNTY HOSPITAL provides quick medical care for emergency cases brought to the hospital by ambulance or private automobile. At any hour there is always one doctor on duty in the emergency room. However, because of a growing tendency for emergency cases to use these facilities rather than go to a private physician, the hospital has been experiencing a continuing increase in the number of emergency room visits each year. As a result, it has become quite common for patients arriving during peak usage hours (the early evening) to have to wait until it is their turn to be treated by the doctor. Therefore, a proposal has been made that a second doctor should be assigned to the emergency room during these hours, so that two emergency cases can be treated simultaneously. The hospital's management engineer has been assigned to study this question.[1]

The management engineer began by gathering the relevant historical data and then projecting these data into the next year. Recognizing that the emergency room is a queueing system, she applied several alternative queueing theory models to predict the waiting characteristics of the system with one doctor and with two doctors, as you will see in the latter sections of this chapter (see Tables 17.2, 17.3, and 17.4).

17.2 BASIC STRUCTURE OF QUEUEING MODELS

The Basic Queueing Process

The basic process assumed by most queueing models is the following. *Customers* requiring service are generated over time by an *input source.* These customers enter the *queueing system* and join a *queue.* At certain times, a member of the queue is selected for service by some rule known as the *queue discipline.* The required service is then performed for the customer by the *service mechanism,* after which the customer leaves the queueing system. This process is depicted in Fig. 17.1.

Many alternative assumptions can be made about the various elements of the queueing process; they are discussed next.

Input Source (Calling Population)

One characteristic of the input source is its size. The *size* is the total number of customers that might require service from time to time, i.e., the total number of distinct potential customers. This population from which arrivals come is referred to as the **calling population.** The size may be assumed to be either *infinite* or *finite* (so that the input source also is said to be either *unlimited* or *limited*). Because the calculations are far easier for the infinite case, this assumption often is made even when the actual size is some rela-

[1]For one actual case study of this kind, see W. Blaker Bolling, "Queueing Model of a Hospital Emergency Room," *Industrial Engineering,* September 1972, pp. 26–31.