Chapter 5
Markov Chains

Many decisions must be made within the context of randomness. Random failures of equipment, fluctuating production rates, and unknown demands are all part of normal decision making processes. In an effort to quantify, understand, and predict the effects of randomness, the mathematical theory of probability and stochastic processes has been developed, and in this chapter, one special type of stochastic process called a Markov chain is introduced. In particular, a Markov chain has the property that the future is independent of the past given the present. These processes are named after the probabilist A. A. Markov who published a series of papers starting in 1907 which laid the theoretical foundations for finite state Markov chains.

An interesting example from the second half of the 19^{th} century is the so-called Galton-Watson process [1]. (Of course, since this was before Markov's time, Galton and Watson did not use Markov chains in their analyses, but the process they studied is a Markov chain and serves as an interesting example.) Galton, a British scientist and first cousin of Charles Darwin, and Watson, a clergyman and mathematician, were interested in answering the question of when and with what probability would a given family name become extinct. In the 19^{th} century, the propagation or extinction of aristocratic family names was important, since land and titles stayed with the name. The process they investigated was as follows: At generation zero, the process starts with a single ancestor. Generation one consists of all the sons of the ancestor (sons were modeled since it was the male that carried the family name). The next generation consists of all the sons of each son from the first generation (i.e., grandsons to the ancestor), generations continuing ad infinitum or until extinction. The assumption is that for each individual in a generation, the probability of having zero, one, two, etc. sons is given by some specified (and unchanging) probability mass function, and that mass function is identical for all individuals at any generation. Such a process might continue to expand or it might become extinct, and Galton and Watson were able to address the questions of whether or not extinction occurred and, if extinction did occur, how many generations would it take. The distinction that makes a Galton-Watson process a Markov chain is the fact that at any generation, the number of individuals in the next generation is completely independent of the number of individuals in previous generations, as long as the

R.M. Feldman, C. Valdez-Flores, *Applied Probability and Stochastic Processes*, 2nd ed., DOI 10.1007/978-3-642-05158-6_5,

number of individuals in the current generation are known. It is processes with this feature (the future being independent of the past given the present) that will be studied next. They are interesting, not only because of their mathematical elegance, but also because of their practical utilization in probabilistic modeling.

5.1 Basic Definitions

In this chapter we study a special type of discrete parameter stochastic process that is one step more complicated than a sequence of *i.i.d.* random variables; namely, Markov chains. Intuitively, a Markov chain is a discrete parameter process in which the future is independent of the past given the present. For example, suppose that we decided to play a game with a fair, unbiased coin. We each start with five pennies and repeatedly toss the coin. If it turns up heads, then you give me a penny; if tails, I give you a penny. We continue until one of us has none and the other has ten pennies. The sequence of heads and tails from the successive tosses of the coin would form an *i.i.d.* stochastic process; the sequence representing the total number of pennies that you hold would be a Markov chain. To see this, assume that after several tosses, you currently have three pennies. The probability that after the next toss you will have four pennies is 0.5 and knowledge of the past (i.e., how many pennies you had one or two tosses ago) does not help in calculating the probability of 0.5. Thus, the future (how many pennies you will have after the next toss) is independent of the past (how many pennies you had several tosses ago) given the present (you currently have three pennies).

Another example of the Markov property comes from Mendelian genetics. Mendel in 1865 demonstrated that the seed color of peas was a genetically controlled trait. Thus, knowledge about the gene pool of the current generation of peas is sufficient information to predict the seed color for the next generation. In fact, if full information about the current generation's genes are known, then knowing about previous generations does not help in predicting the future; thus, we would say that the future is independent of the past given the present.

Definition 5.1. The stochastic process $X = \{X_n; n = 0, 1, \cdots\}$ with discrete state space E is a *Markov chain* if the following holds for each $j \in E$ and $n = 0, 1, \cdots$

$$\Pr\{X_{n+1} = j | X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n\} = \Pr\{X_{n+1} = j | X_n = i_n\} ,$$

for any set of states $i_0, \cdots, i_n$ in the state space. Furthermore, the Markov chain is said to have *stationary* transition probabilities if

$$\Pr\{X_1 = j | X_0 = i\} = \Pr\{X_{n+1} = j | X_n = i\} .$$

□

The first equation in Definition 5.1 is a mathematical statement of the Markov property. To interpret the equation, think of time n as the present. The left-hand-side

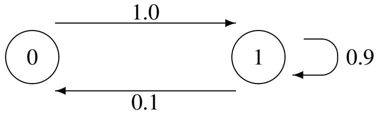


Fig. 5.1 State diagram for the Markov chain of Example 5.1

is the probability of going to state j next, given the history of all past states. The right-hand-side is the probability of going to state j next, given only the present state. Because they are equal, we have that the past history of states provides no additional information helpful in predicting the future if the present state is known. The second equation (i.e., the stationary property) simply indicates that the probability of a one-step transition does not change as time increases (in other words, the probabilities are the same in the winter and the summer).

In this chapter it is always assumed that we are working with stationary transition probabilities. Because the probabilities are stationary, the only information needed to describe the process are the initial conditions (a probability mass function for X_0) and the one-step transition probabilities. A square matrix is used for the transition probabilities and is often denoted by the capital letter $\mathbf{P}$, where

$$P(i,j) = \Pr\{X_1 = j | X_0 = i\} \ . \tag{5.1}$$

Since the matrix $\mathbf{P}$ contains probabilities, it is always nonnegative (a matrix is nonnegative if every element of it is nonnegative) and the sum of the elements within each row equals one. In fact, any nonnegative matrix with row sums equal to one is called a *Markov matrix*.

Example 5.1. Consider a farmer using an old tractor. The tractor is often in the repair shop but it always takes only one day to get it running again. The first day out of the shop it always works but on any given day thereafter, independent of its previous history, there is a 10% chance of it not working and thus being sent back to the shop. Let $X_0, X_1, \cdots$ be random variables denoting the daily condition of the tractor, where a one denotes the working condition and a zero denotes the failed condition. In other words, $X_n = 1$ denotes that the tractor is working on day n and $X_n = 0$ denotes it being in the repair shop on day n. Thus, $X_0, X_1, \cdots$ is a Markov chain with state space $E = \{0, 1\}$ and with Markov matrix

$$\mathbf{P} = \begin{array}{c} 0 \\ 1 \end{array} \begin{bmatrix} 0 & 1 \\ 0.1 & 0.9 \end{bmatrix} .$$

(Notice that the state space is sometimes written to the left of the matrix to help in keeping track of which row refers to which state.) □

In order to develop a mental image of the Markov chain, it is very helpful to draw a state diagram (Fig. 5.1) of the Markov matrix. In the diagram, each state is represented by a circle and the transitions with positive probabilities are represented by an arrow. Until the student is very familiar with Markov chains, we recommend that state diagrams be drawn for any chain that being discussed.

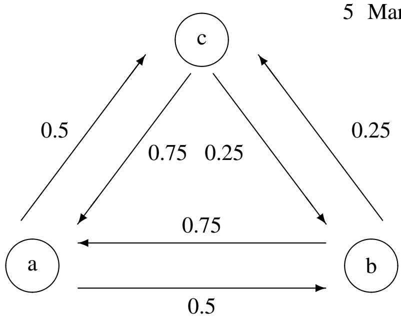


Fig. 5.2 State diagram for the Markov chain of Example 5.2

Example 5.2. A salesman lives in town 'a' and is responsible for towns 'a', 'b', and 'c'. Each week he is required to visit a different town. When he is in his home town, it makes no difference which town he visits next so he flips a coin and if it is heads he goes to 'b' and if tails he goes to 'c'. However, after spending a week away from home he has a slight preference for going home so when he is in either towns 'b' or 'c' he flips two coins. If two heads occur, then he goes to the other town; otherwise he goes to 'a'. The successive towns that he visits form a Markov chain with state space $E = \{a, b, c\}$ where the random variable X_n equals a, b, or c according to his location during week n. The state diagram for this system is given in Fig. 5.2 and the associated Markov matrix is

$$\mathbf{P} = \begin{array}{c} a \\ b \\ c \end{array} \begin{bmatrix} 0 & 0.50 & 0.50 \\ 0.75 & 0 & 0.25 \\ 0.75 & 0.25 & 0 \end{bmatrix} .$$

□

Example 5.3. Let $X = \{X_n; n = 0, 1, \cdots\}$ be a Markov chain with state space $E = \{1, 2, 3, 4\}$ and transition probabilities given by

$$\mathbf{P} = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.3 & 0.7 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0.2 & 0 & 0.1 & 0.7 \end{bmatrix} .$$

The chain in this example is structurally different than the other two examples in that you might start in State 4, then go to State 3, and never get to State 1; or you might start in State 4 and go to State 1 and stay there forever. The other two examples, however, involved transitions in which it was always possible to eventually reach every state from every other state. □

Example 5.4. The final example in this section is taken from Parzen [3, p. 191] and illustrates that the parameter n need not refer to time. (It is often true that the "steps" of a Markov chain refer to days, weeks, or months, but that need not be the case.) Consider a page of text and represent vowels by zeros and consonants by ones.

Fig. 5.3 State diagram for the Markov chain of Example 5.3

Thus the page becomes a string of zeros and ones. It has been indicated that the sequence of vowels and consonants in the Samoan language forms a Markov chain, where a vowel always follows a consonant and a vowel follows another vowel with a probability of 0.51 [2]. Thus, the sequence of ones and zeros on a page of Samoan text would evolve according to the Markov matrix

$$\mathbf{P} = \begin{array}{c} 0 \\ 1 \end{array} \begin{bmatrix} 0.51 & 0.49 \\ 1 & 0 \end{bmatrix} .$$

□

After a Markov chain has been formulated, there are many questions that might be of interest. For example: In what state will the Markov chain be five steps from now? What percent of time is spent in a given state? Starting from one state, will the chain ever reach another fixed state? If a profit is realized for each visit to a particular state, what is the long run average profit per unit of time? The remainder of this chapter is devoted to answering questions of this nature.

- *Suggestion: Do Part (a) of Problems 5.1–5.3 and 5.6–5.9.*

5.2 Multistep Transitions

The Markov matrix provides direct information about one-step transition probabilities and it can also be used in the calculation of probabilities for transitions involving more than one step. Consider the salesman of Example 5.2 starting in town b. The Markov matrix indicates that the probability of being in State a after one step (in one week) is 0.75, but what is the probability that he will be in State a after two steps? Figure 5.4 illustrates the paths that go from b to a in two steps (some of the paths shown have probability zero). Thus, to compute the probability, we need to sum over all possible routes. In other words we would have the following calculations:

$$\begin{aligned} \Pr\{X_2 = a | X_0 = b\} &= \Pr\{X_1 = a | X_0 = b\} \times \Pr\{X_2 = a | X_1 = a\} \\ &\quad + \Pr\{X_1 = b | X_0 = b\} \times \Pr\{X_2 = a | X_1 = b\} \\ &\quad + \Pr\{X_1 = c | X_0 = b\} \times \Pr\{X_2 = a | X_1 = c\} \\ &= P(b,a)P(a,a) + P(b,b)P(b,a) + P(b,c)P(c,a) . \end{aligned}$$

The final equation should be recognized as the definition of matrix multiplication; thus,

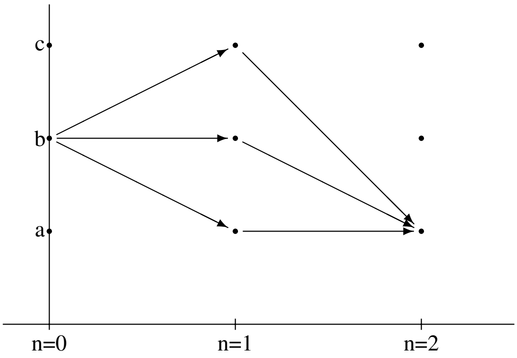


Fig. 5.4 Possible paths of a two-step transition from State b to State a for a three-state Markov chain

$$\Pr\{X_2 = a | X_0 = b\} = P^2(b,a) .$$

This result is easily generalized into the following property:

Property 5.1. *Let $X = \{X_n; n = 0, 1, \cdots\}$ be a Markov chain with state space E and Markov matrix $\mathbf{P}$, then for $i, j \in E$ and $n = 1, 2, \cdots$*

$$\Pr\{X_n = j | X_0 = i\} = P^n(i,j) ,$$

where the right-hand side represents the $i-j$ element of the matrix $\mathbf{P}^n$.

Property 5.1 indicates that $P^n(i,j)$ is interpreted to be the probability of going from state i to state j in n steps. It is important to remember that the notation $P^n(i,j)$ means that the matrix is *first* raised to the n^{th} power and then the $i-j$ element of the resulting matrix is taken for the answer.

Returning to Example 5.2, the squared matrix is

$$\mathbf{P}^2 = \begin{bmatrix} 0.75 & 0.125 & 0.125 \\ 0.1875 & 0.4375 & 0.375 \\ 0.1875 & 0.375 & 0.4375 \end{bmatrix} .$$

The $b-a$ element of $\mathbf{P}^2$ is 0.1875 and thus there is a 0.1875 probability of being in town a two weeks after being in town b.

Markov chains are often used to analyze the cost or profit of an operation and thus we need to consider a cost or profit function imposed on the process. For example, suppose in Example 5.2 that every week spent in town a resulted in a profit of \$1000, every week spent in town b resulted in a profit of \$1200, and every week spent in town c resulted in a profit of \$1250. We then might ask what would be the expected profit after the first week if the initial town was town a? Or, more generally, what would be the expected profit of the n^{th} week if the initial town was a? It should not be too difficult to convince yourself that the following calculation would be appropriate:

$$E[\text{Profit for week } n] = P^n(a,a) \times 1000 + P^n(a,b) \times 1200 + P^n(a,c) \times 1250 \,.$$

We thus have the following property.

Property 5.2. *Let $X = \{X_n; n = 0, 1, \cdots\}$ be a Markov chain with state space E, Markov matrix* $\mathbf{P}$*, and profit function* $\mathbf{f}$ *(i.e., each time the chain visits state i, a profit of $f(i)$ is obtained). The expected profit at the n^{th} step is given by*

$$E[f(X_n)|X_0 = i] = \mathbf{P}^n\mathbf{f}(i) \,.$$

Note again that in the right-hand side of the equation, the matrix $\mathbf{P}^n$ is multiplied by the vector $\mathbf{f}$ first and then the i^{th} component of the resulting vector is taken as the answer. Thus, in Example 5.2, we have that the expected profit during the second week given that the initial town was a is

$$\begin{aligned} E[f(X_2)|X_0 = a] &= 0.75 \times 1000 + 0.125 \times 1200 + 0.125 \times 1250 \\ &= 1056.25. \end{aligned}$$

Up until now we have always assumed that the initial state was known. However, that is not always the situation. The manager of the traveling salesman might not know for sure the location of the salesman; instead, all that is known is a probability mass function describing his initial location. Again, using Example 5.2, suppose that we do not know for sure the salesman's initial location but know that there is a 50% chance he is in town a, a 30% chance he is in town b, and a 20% chance he is in town c. We now ask, what is the probability that the salesman will be in town a next week? The calculations for that are

$$\Pr\{X_1 = a\} = 0.50 \times 0 + 0.30 \times 0.75 + 0.20 \times 0.75 = 0.375 \,.$$

These calculations generalize to the following.

Property 5.3. *Let $X = \{X_n; n = 0, 1, \cdots\}$ be a Markov chain with state space E, Markov matrix* $\mathbf{P}$*, and initial probability vector* $\boldsymbol{\mu}$ *(i.e., $\mu(i) = \Pr\{X_0 = i\}$). Then*

$$\Pr_\mu\{X_n = j\} = \boldsymbol{\mu}\mathbf{P}^n(j) \,.$$

It should be noticed that μ is a subscript to the probability statement on the left-hand side of the equation. The purpose for the subscript is to insure that there is no confusion over the given conditions. And again when interpreting the right-hand-side of the equation, the vector $\boldsymbol{\mu}$ is multiplied by the matrix $\mathbf{P}^n$ first and then the j^{th} element is taken from the resulting vector.

The last two properties can be combined, when necessary, into one statement.

Property 5.4. *Let $X = \{X_n; n = 0, 1, \cdots\}$ be a Markov chain with state space E, Markov matrix* **P***, initial probability vector* $\boldsymbol{\mu}$*, and profit function* **f**. *The expected profit at the n^{th} step is given by*

$$E_\mu[f(X_n)] = \boldsymbol{\mu}\mathbf{P}^n\mathbf{f}\,.$$

Returning to Example 5.2 and using the initial probabilities and profit function given above, we have that the expected profit in the second week is calculated to be

$$\begin{aligned}\boldsymbol{\mu}\mathbf{P}^2\mathbf{f} &= (0.50, 0.30, 0.20)\begin{bmatrix} 0.75 & 0.125 & 0.125 \\ 0.1875 & 0.4375 & 0.375 \\ 0.1875 & 0.375 & 0.4375 \end{bmatrix}\begin{pmatrix} 1000 \\ 1200 \\ 1250 \end{pmatrix} \\ &= 1119.375\,.\end{aligned}$$

Example 5.5. A market analysis concerning consumer behavior in auto purchases has been conducted. Body styles traded-in and purchased have been recorded by a particular dealer with the following results:

Number of Customers	Trade
275	sedan for sedan
180	sedan for station wagon
45	sedan for convertible
80	station wagon for sedan
120	station wagon for station wagon
150	convertible for sedan
50	convertible for convertible

These data are believed to be representative of average consumer behavior, and it is assumed that the Markov assumptions are appropriate.

We shall develop a Markov chain to describe the changing body styles over the life of a customer. (Notice, for purposes of instruction, we are simplifying the process by assuming that the age of the customer does not affect the customer's choice of a body style.) We define the Markov chain to be the body-style of the automobile that the customer has immediately after a trade-in, where the state space is $E = \{s, w, c\}$ with s for the sedan, w for the wagon, and c for the convertible. Of 500 customers who have a sedan, 275 will stay with the sedan during the next trade; therefore, the s-s element of the transition probability matrix will be 275/500. Thus, the Markov matrix for the "body style" Markov chain is

$$\mathbf{P} = \begin{array}{c} s \\ w \\ c \end{array}\begin{bmatrix} 275/500 & 180/500 & 45/500 \\ 80/200 & 120/200 & 0 \\ 150/200 & 0 & 50/200 \end{bmatrix} = \begin{bmatrix} 0.55 & 0.36 & 0.09 \\ 0.40 & 0.60 & 0.0 \\ 0.75 & 0.0 & 0.25 \end{bmatrix}.$$

Let us assume we have a customer whose behavior is described by this model. Furthermore, the customer always buys a new car in January of every year. It is now January 1997, and the customer enters the dealership with a sedan (i.e., $X_{1996} = s$). From the above matrix, $\Pr\{X_{1997} = s | X_{1996} = s\} = 0.55$, or the probability that the customer will leave with another sedan is 55%. We would now like to predict what body-style the customer will have for trade-in during January 2000. Notice that the question involves three transitions of the Markov chain; therefore, the cubed matrix must be determined, which is

$$\mathbf{P}^3 = \begin{bmatrix} 0.5023 & 0.4334 & 0.0643 \\ 0.4816 & 0.4680 & 0.0504 \\ 0.5355 & 0.3780 & 0.0865 \end{bmatrix} .$$

Thus, there is approximately a 50% chance the customer will enter in the year 2000 with a sedan, a 43% chance with a station wagon, and almost a 7% chance of having a convertible. The probability that the customer will leave this year (1997) with a convertible and then leave next year (1998) with a sedan is given by $P(s,c) \times P(c,s) = 0.09 \times 0.75 = 0.0675$. Now determine the probability that the customer who enters the dealership now (1997) with a sedan will leave with a sedan and also leave with a sedan in the year 2000. The mathematical statement is

$$\Pr\{X_{1997} = s, X_{2000} = s | X_{1996} = s\} = P(s,s) \times P^3(s,s) \approx 0.28 .$$

Notice in the above probability statement, that no mention is made of the body style for the intervening years; thus, the customer may or may not switch in the years 1998 and 1999.

Now to illustrate profits. Assume that a sedan yields a profit of \$1200, a station wagon yields \$1500, and a convertible yields \$2500. The expected profit this year from the customer entering with a sedan is \$1425, and the expected profit in the year 1999 from a customer who enters the dealership with a sedan in 1997 is approximately \$1414. Or, to state this mathematically,

$$E[f(X_{1999}) | X_{1996} = s] \approx 1414 ,$$

where $\mathbf{f} = (1200, 1500, 2500)^T$. □

- *Suggestion: Do Problems 5.10a–c, 5.11a–f, and 5.12.*

5.3 Classification of States

There are two types of states possible in a Markov chain, and before most questions can be answered, the individual states for a particular chain must be classified into one of these two types. As we begin the Markov chain study, it is important to cover a significant amount of new notation. The student will discover that the time spent in learning the new terminology will be rewarded with a fuller understanding of

Markov chains. Furthermore, a good understanding of the dynamics involved for Markov chains will greatly aid in the understanding of the entire area of stochastic processes.

Two random variables that will be extremely important denote "first passage times" and "number of visits to a fixed state". To describe these random variables, consider a fixed state, call it State j, in the state space for a Markov chain. The first passage time is a random variable, denoted by T^j, that equals the time (i.e., number of steps) it takes to reach the fixed state *for the first time*. Mathematically, the first passage time random variable is defined by

$$T^j = \min\{n \geq 1 : X_n = j\} , \tag{5.2}$$

where the minimum of the empty set is taken to be $+\infty$. For example, in Example 5.3, if $X_0 = 1$ then $T^2 = \infty$, i.e., if the chain starts in State 1 then it will never reach State 2.

The number of visits to a state is a random variable, denoted by N^j, that equals the total number of visits (including time zero) that the Markov chain makes to the fixed state throughout the life of the chain. Mathematically, the "number of visits" random variable is defined by

$$N^j = \sum_{n=0}^{\infty} I(X_n, j) , \tag{5.3}$$

where I is the identity matrix. The identity matrix is used simply as a "counter". Because the identity has ones on the diagonal and zeroes off the diagonal, it follows that $I(X_n, j) = 1$ if $X_n = j$; otherwise, it is zero. It should be noted that the summation in (5.3) starts at $n = 0$; thus, if $X_0 = j$ then N^j must be at least one.

Example 5.6. Let us consider a realization for the Markov chain of Example 5.3. (By realization, we mean conceptually that an experiment is conducted and we record the random outcomes of the chain.) Assume that the first part of the realization (Fig. 5.5) is $X_0 = 4, X_1 = 4, X_2 = 4, X_3 = 3, X_4 = 2, X_5 = 3, X_6 = 2, \cdots$. The first passage random variables for this realization are $T^1 = \infty, T^2 = 4, T^3 = 3,$ and $T^4 = 1$. To see why $T^1 = \infty$, it is easiest to refer to the state diagram describing the Markov chain (Fig. 5.3). By inspecting the diagram it becomes obvious that once the chain is in either States 2 or 3 that it will never get to State 1 and thus $T^1 = \infty$.

Using the same realization, the number of visits random variables are $N^1 = 0, N^2 = \infty, N^3 = \infty$, and $N^4 = 3$. Again, the values for N^1, N^2, and N^3 are obtained by inspecting the state diagram and observing that if the chain ever gets to States 2 or 3 it will stay in those two states forever and thus will visit them an infinite number of times.

Let us perform the experiment one more time. Assume that our second realization results in the values $X_0 = 4, X_1 = 4, X_2 = 1, X_3 = 1, X_4 = 1, \cdots$. The new outcomes for the first passage random variables for this second realization are $T^1 = 2, T^2 = \infty, T^3 = \infty,$ and $T^4 = 1$. Furthermore, $N^1 = \infty, N^2 = 0, N^3 = 0,$ and $N^4 = 2$. Thus, you should understand from this example that we may not be able to say what the value

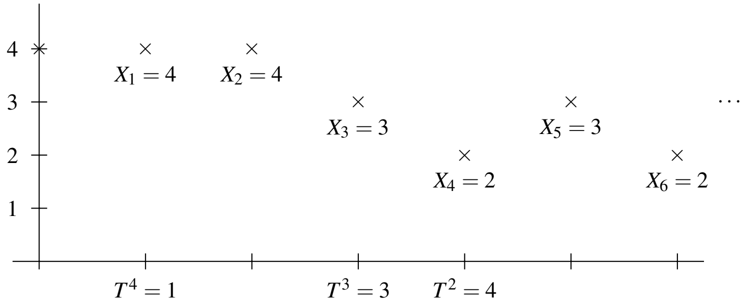


Fig. 5.5 One Possible Realization for the Markov chain of Example5.1

of T^j will be before an experiment, but we should be able to describe its probability mass function. □

The primary quantity of interest regarding the first passage times are the so-called first passage probabilities. The major question of interest is whether or not it is possible to reach a particular state from a given initial state. To answer this question, we determine the first passage probability, denoted by $F(i,j)$, which is the probability of eventually reaching State j at least once given that the initial state was i. The probability $F(i,j)$ ("F" for *first* passage) is defined by

$$F(i,j) = \Pr\{T^j < \infty | X_0 = i\} \,. \tag{5.4}$$

By inspecting the state diagram in Fig. 5.3, it should be obvious that the first passage probabilities for the chain of Example 5.3 are given by

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ <1 & <1 & <1 & <1 \end{bmatrix} .$$

The primary quantity of interest for the number of visits random variable is its expected value. The expected number of visits to State j given the initial state was i is denoted by $R(i,j)$ ("R" for *returns*) and is defined by

$$R(i,j) = E[N^j | X_0 = i] \,. \tag{5.5}$$

Again the state diagram of Fig. 5.3 allows the determination of some of the values of R as follows:

$$\mathbf{R} = \begin{bmatrix} \infty & 0 & 0 & 0 \\ 0 & \infty & \infty & 0 \\ 0 & \infty & \infty & 0 \\ \infty & \infty & \infty & <\infty \end{bmatrix} .$$

The above matrix may appear unusual because of the occurrence of ∞ for elements of the matrix. In Sect. 5.5, numerical methods will be given to calculate the values of the **R** and **F** matrices; however, the values of ∞ are obtained by an understanding of the structures of the Markov chain and not through specific formulas. For now, our goal is to develop an intuitive understanding of these processes; therefore, the major concern for the **F** matrix is whether an element is zero or one, and the concern for the **R** matrix is whether an element is zero or infinity. As might be expected, there is a close relation between $R(i,j)$ and $F(i,j)$ as is shown in the following property.

Property 5.5. *Let $R(i,j)$ and $F(i,j)$ be as defined in (5.5) and (5.4), respectively. Then*

$$R(i,j) = \begin{cases} \frac{1}{1-F(j,j)} \text{ for } i = j, \\ \\ \frac{F(i,j)}{1-F(j,j)} \text{ for } i \neq j; \end{cases},$$

where the convention $0/0 = 0$ is used.

The above discussion utilizing Example 5.3 should help to point out that there is a basic difference between States 1, 2 or 3 and State 4 of the example. Consider that if the chain ever gets to State 1 or to State 2 or to State 3 then that state will continually reoccur. However, even if the chain starts in State 4, it will only stay there a finite number of times and will eventually leave State 4 never to return. These ideas give rise to the terminology of recurrent and transient states. Intuitively, a state is called recurrent if, starting in that state, it will continuously reoccur; and a state is called transient if, starting in that state, it will eventually leave that state never to return. Or equivalently, a state is recurrent if, starting in that state, the chain must (i.e., with probability one) eventually return to the state at least once; and a state is transient if, starting in that state, there is a chance (i.e., with probability greater than zero) that the chain will leave the state and never return. The mathematical definitions are based on these notions developed above.

Definition 5.2. A state j is called *transient* if $F(j,j) < 1$. Equivalently, state j is *transient* if $R(j,j) < \infty$. □

Definition 5.3. A state j is called *recurrent* if $F(j,j) = 1$. Equivalently, state j is *recurrent* if $R(j,j) = \infty$. □

From the above two definitions, a state must either be transient or recurrent. A dictionary[1] definition for the word *transient* is "passing esp. quickly into and out of existence". Thus, the use of the word is justified since transient states will only occur for a finite period of time. For a transient state, there will be a time after which the transient state will never again be visited. A dictionary[1] definition for *recurrent* is "returning or happening time after time". So again the mathematical concept of a

[1] *Webster's Ninth New Collegiate Dictionary*, (Springfield, MA: Merriam-Webster, Inc., 1989)

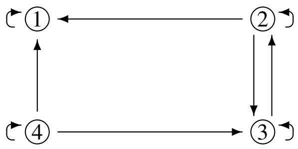


Fig. 5.6 State diagram for the Markov chain of Example 5.7

recurrent state parallels the common English usage: recurrent states are recognized by those states which are continually revisited.

Recurrent states might also be periodic, as with the Markov matrix

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

but discussions regarding periodic chains will be left to other texts.

By drawing the state diagrams for the Markov chains of Examples 5.1 and 5.2, it should become clear that all states in those two cases are recurrent. The only transient state so far illustrated is State 4 of Example 5.3.

Example 5.7. Let $X = \{X_n; n = 0, 1, \cdots\}$ be a Markov chain with state space $E = \{1, 2, 3, 4\}$ and transition probabilities given by

$$\mathbf{P} = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.01 & 0.29 & 0.7 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0.2 & 0 & 0.1 & 0.7 \end{bmatrix}.$$

Notice the similarity between this Markov chain and the chain in Example 5.3. Although the numerical difference between the two Markov matrices is slight, there is a radical difference between the structures of the two chains. (The difference between a zero term and a nonzero term, no matter how small, can be very significant.) The chain of this example has one recurrent state and three transient states.

By inspecting the state diagram for this Markov chain (Fig. 5.6) the following matrices are obtained (through observation, not through calculation):

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & <1 & <1 & 0 \\ 1 & 1 & <1 & 0 \\ 1 & <1 & <1 & <1 \end{bmatrix}.$$

$$\mathbf{R} = \begin{bmatrix} \infty & 0 & 0 & 0 \\ \infty & <\infty & <\infty & 0 \\ \infty & <\infty & <\infty & 0 \\ \infty & <\infty & <\infty & <\infty \end{bmatrix}.$$

It should be observed that $F(3,2) = 1$ even though both State 2 and State 3 are transient. It is only the diagonal elements of the **F** and **R** matrices that determine whether a state is transient or recurrent. □

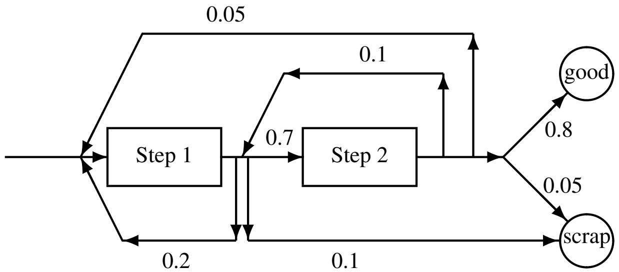


Fig. 5.7 Two step manufacturing process of Example 5.8

Example 5.8. A manufacturing process consists of two processing steps in sequence. After step 1, 20% of the parts must be reworked, 10% of the parts are scrapped, and 70% proceed to the next step. After step 2, 5% of the parts must be returned to the first step, 10% must be reworked and 5% are scrapped; the remainder are sold. The diagram of Fig. 5.7 illustrates the dynamics of the manufacturing process.

The Markov matrix associated with this manufacturing process is given by

$$\mathbf{P} = \begin{array}{c} 1 \\ 2 \\ s \\ g \end{array} \begin{bmatrix} 0.2 & 0.7 & 0.1 & 0.0 \\ 0.05 & 0.1 & 0.05 & 0.8 \\ 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix} . \tag{5.6}$$

Consider the dynamics of the chain. A part to be manufactured will begin the process by entering State 1. After possibly cycling for awhile, the part will end the process by entering either the "g" state or the "s" state. Therefore, States 1 and 2 are transient, and States g and s are recurrent. □

Along with classifying states, we also need to be able to classify sets of states. We will first define a closed set which is a set that once the Markov chain has entered the set, it cannot leave the set.

Definition 5.4. Let $X = \{X_n; n = 0, 1, \cdots\}$ be a Markov chain with Markov matrix $\mathbf{P}$ and let C be a set of states contained in its state space. Then C is *closed* if

$$\sum_{j \in C} P(i,j) = 1 \text{ for all } i \in C .$$

□

To illustrate the concept of closed sets, refer again to Example 5.3. The sets $\{2,3,4\}$, $\{2\}$, and $\{3,4\}$ are *not* closed sets. The set $\{1,2,3,4\}$ is obviously closed, but it can be reduced to a smaller closed set. The set $\{1,2,3\}$ is also closed, but again it can be further reduced. Both sets $\{1\}$ and $\{2,3\}$ are closed and cannot be reduced further. This idea of taking a closed set and trying to reduce it is extremely important and leads to the definition of an irreducible set.

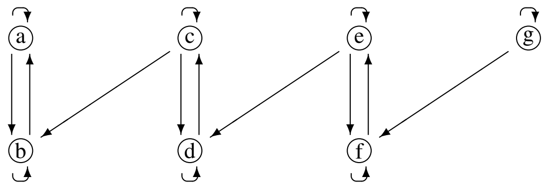


Fig. 5.8 State diagram for the Markov chain of Example 5.9

Definition 5.5. A closed set of states that contains no proper subset which is also closed is called *irreducible*. A state that forms an irreducible set by itself is called an *absorbing* state. □

The Markov chain of Example 5.3 has two irreducible sets: the set $\{1\}$ and the set $\{2,3\}$. Since the first irreducible set contains only one state, that state is an absorbing state.

Example 5.9. Let X be a Markov chain with state space $E = \{a, \cdots, g\}$ and Markov matrix

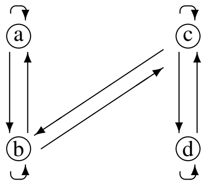
$$\mathbf{P} = \begin{array}{c} a \\ b \\ c \\ d \\ e \\ f \\ g \end{array} \begin{bmatrix} 0.3 & 0.7 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.2 & 0.4 & 0.4 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.8 & 0.1 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 0.7 & 0.3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.4 & 0.6 \end{bmatrix}.$$

By drawing a state diagram (Fig. 5.8), you should observe that states a and b are recurrent and all others are transient. (If that is not obvious, notice that once the chain reaches either a or b, it will stay in those two states, there are no paths leading away from the set $\{a,b\}$; furthermore, all states will eventually reach State b). Obviously, the entire state space is closed. Since no state goes to g, we can exclude g and still have a closed set; namely, the set $\{a,b,c,d,e,f\}$ is closed. However, the set $\{a,b,c,d,e,f\}$ can be reduced to a smaller closed set so it is *not* irreducible. Because there is a path going from e to f, the set $\{a,b,c,d,e\}$ is *not* closed. By excluding the state e, a closed set is again obtained; namely, $\{a,b,c,d\}$ is also closed since there are no paths out of the set $\{a,b,c,d\}$ to another state. Again, however, it can also be reduced so the set $\{a,b,c,d\}$ is *not* irreducible. Finally, consider the set $\{a,b\}$. This two-state set is closed and cannot be further reduced to a smaller closed set; therefore, the set $\{a,b\}$ is an irreducible set. □

The reason that Definitions 5.4 and 5.5 are important is the following property.

Property 5.6. *All states within an irreducible set are of the same classification.*

Fig. 5.9 State diagram for the Markov chain of Eq. 5.7



The significance of the above property is that if you can identify one state within an irreducible set as being transient, then all states within the set are transient, and if one state is recurrent, then all states within the set are recurrent. We are helped even more by recognizing that it is impossible to have an irreducible set of transient states if the set contains only a finite number of states; thus we have the following property.

Property 5.7. *Let C be an irreducible set of states such that the number of states within C is finite. Then each state within C is recurrent.*

There is one final concept that will help in identifying irreducible sets; namely, communication between states. Communication between states is like communication between people; there is communication only if messages can go both ways. In other words, two states, i and j, *communicate* if and only if it is possible to eventually reach j from i *and* it is possible to eventually reach i from j. In Example 5.3, States 2 and 3 communicate, but States 4 and 2 do not communicate. Although State 2 can be reached from 4, it does not go both ways because State 4 cannot be reached from 2. The communication must be both ways but it does not have to be in one step. For example, in the Markov chain with state space $\{a,b,c,d\}$ and with Markov matrix

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.2 & 0.4 & 0.4 & 0 \\ 0 & 0.1 & 0.8 & 0.1 \\ 0 & 0 & 0.3 & 0.7 \end{bmatrix}, \tag{5.7}$$

all states communicate with every other state. And, in particular, State a communicates with State d even though they can not reach each other in one step (see Fig. 5.9).

The notion of communication is often the concept used to identify irreducible sets as is given in the following property.

Property 5.8. *The closed set of states C is irreducible if and only if every state within C communicates with every other state within C.*

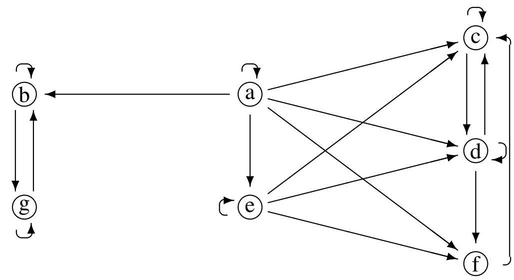


Fig. 5.10 State diagram for the Markov chain of Example 5.3

The procedure for identifying irreducible sets of states for a Markov chain with a finite state space is to first draw the state diagram, then pick a state and identify all states that communicate with it. If the set made up of all those states that communicate with it is closed, then the set is an irreducible, recurrent set of states. If it is not closed, then the originally chosen state and all the states that communicate with it are transient states.

Example 5.10. Let X be a Markov chain with state space $E = \{a, \cdots, g\}$ and Markov matrix

$$\mathbf{P} = \begin{array}{c} a \\ b \\ c \\ d \\ e \\ f \\ g \end{array} \begin{bmatrix} 0.3 & 0.1 & 0.2 & 0.2 & 0.1 & 0.1 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0.4 & 0.6 & 0 & 0 & 0 \\ 0 & 0 & 0.3 & 0.2 & 0 & 0.5 & 0 \\ 0 & 0 & 0.2 & 0.3 & 0.4 & 0.1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 & 0 & 0 & 0.2 \end{bmatrix} .$$

By drawing a state diagram (Fig. 5.10), it is seen that there are two irreducible recurrent sets of states which are $\{b, g\}$ and $\{c, d, f\}$. States a and e are transient.

□

- *Suggestion: Do Problems 5.13a–c and 5.14a–d.*

5.4 Steady-State Behavior

The reason that so much effort has been spent in classifying states and sets of states is that the analysis of the limiting behavior of a chain is dependent on the type of state under consideration. To illustrate long-run behavior, we again return to Example 5.2 and focus attention on State a. Figure 5.11 shows a graph which gives the probabilities of being in State a at time n given that at time zero the chain was in State a. (In other words, the graph gives the values of $P^n(a,a)$ as a function of n.) Although the graph varies dramatically for small n, it reaches an essentially constant value as n becomes large. It is seen from the graph that

$$\lim_{n\to\infty} \Pr\{X_n = a | X_0 = a\} = 0.42857 .$$

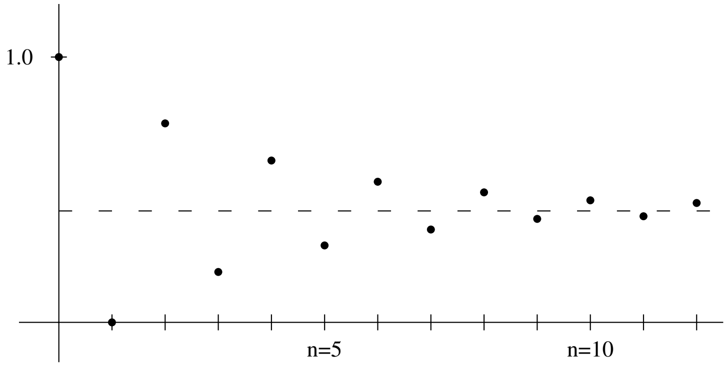


Fig. 5.11 Probabilities from Example 5.2 of being in State a as a function of time

In fact, if you spent the time to graph the probabilities of being in State a starting from State b instead of State a, you would discover the same limiting value, namely,

$$\lim_{n\to\infty} \Pr\{X_n = a | X_0 = b\} = 0.42857 .$$

When discussing steady-state (or limiting) conditions, this is what is meant, not that the chain stops changing — it is dynamic by definition — but that enough time has elapsed so that the probabilities do not change with respect to time. It is often stated that steady-state results are independent of initial conditions, and this is true for the chain in Example 5.2; however it is not always true. In Example 5.3, it is clear that the steady-state conditions are radically different when the chain starts in State 1 as compared to starting in State 2.

If the entire state space of a Markov chain forms an irreducible recurrent set, the Markov chain is called an irreducible recurrent Markov chain. In this case the steady-state probabilities are independent of the initial state and are not difficult to compute as is seen in the following property.

Property 5.9. *Let $X = \{X_n; n = 0, 1, \cdots\}$ be a Markov chain with finite state space E and Markov matrix* **P**. *Furthermore, assume that the entire state space forms an irreducible, recurrent set, and let*

$$\pi(j) = \lim_{n\to\infty} \Pr\{X_n = j | X_0 = i\} .$$

The vector $\boldsymbol{\pi}$ *is the solution to the following system*

$$\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi},$$
$$\sum_{i\in E} \pi(i) = 1 .$$

To illustrate the determination of $\boldsymbol{\pi}$, observe that the Markov chain of Example 5.2 is irreducible, recurrent. Applying Property 5.9, we obtain

$$\begin{aligned}
0.75\pi_b + 0.75\pi_c &= \pi_a \\
0.5\pi_a \qquad\qquad + 0.25\pi_c &= \pi_b \\
0.5\pi_a + 0.25\pi_b \qquad\qquad &= \pi_c \\
\pi_a + \quad \pi_b + \quad \pi_c &= 1 \,.
\end{aligned}$$

There are four equations and only three variables, so normally there would not be a unique solution; however, for an irreducible Markov matrix there is always exactly one redundant equation from the system formed by $\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$. Thus, to solve the above system, arbitrarily choose one of the first three equations to discard and solve the remaining 3 by 3 system (*never* discard the final or norming equation) which yields

$$\pi_a = \frac{3}{7},\ \pi_b = \frac{2}{7},\ \pi_c = \frac{2}{7}\,.$$

Property 5.9 cannot be directly applied to the chain of Example 5.3 because the state space is not irreducible. All irreducible, recurrent sets must be identified and grouped together and then the Markov matrix is rearranged so that the irreducible sets are together and transient states are last. In such a manner, the Markov matrix for a chain can always be rewritten in the form

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & & & & \\ & \mathbf{P}_2 & & & \\ & & \mathbf{P}_3 & & \\ & & & \ddots & \\ \mathbf{B}_1 & \mathbf{B}_2 & \mathbf{B}_3 & \cdots & \mathbf{Q} \end{bmatrix}\,. \tag{5.8}$$

After a chain is in this form, each submatrix $\mathbf{P}_i$ is a Markov matrix and can be considered as an independent Markov chain for which Property 5.9 is applied.

The Markov matrix of Example 5.3 is already in the form of (5.8). Since State 1 is absorbing (i.e., an irreducible set of one state), its associated steady-state probability is easy; namely, $\pi_1 = 1$. States 2 and 3 form an irreducible set so Property 5.9 can be applied to the submatrix from those states resulting in the following system:

$$\begin{aligned}
0.3\pi_2 + 0.5\pi_3 &= \pi_2 \\
0.7\pi_2 + 0.5\pi_3 &= \pi_3 \\
\pi_2 + \quad \pi_3 &= 1
\end{aligned}$$

which yields (after discarding one of the first two equations)

$$\pi_2 = \frac{5}{12} \text{ and } \pi_3 = \frac{7}{12}\,.$$

The values π_2 and π_3 are interpreted to mean that if a snapshot of the chain is taken a long time after it started and if it started in States 2 or 3, then there is a $5/12$

probability that the picture will show the chain in State 2 and a $7/12$ probability that it will be in State 3. Another interpretation of the steady-state probabilities is that if we recorded the time spent in States 2 and 3, then over the long run the fraction of time spent in State 2 would equal $5/12$ and the fraction of time spent in State 3 would equal $7/12$. These steady-state results are summarized in the following property.

Property 5.10. *Let $X = \{X_n; n = 0, 1, \cdots\}$ be a Markov chain with finite state space with k distinct irreducible sets. Let the ℓ^{th} irreducible set be denoted by C_ℓ, and let $\mathbf{P}_\ell$ be the Markov matrix restricted to the ℓ^{th} irreducible set (as in Eq. 5.8). Finally, let $\mathbf{F}$ be the matrix of first passage probabilities.*

- *If State j is transient,*

$$\lim_{n\to\infty} \Pr\{X_n = j | X_0 = i\} = 0\ .$$

- *If State i and j both belong to the ℓ^{th} irreducible set,*

$$\lim_{n\to\infty} \Pr\{X_n = j | X_0 = i\} = \pi(j)$$

  where

$$\boldsymbol{\pi}\mathbf{P}_\ell = \boldsymbol{\pi} \text{ and}$$
$$\sum_{i\in C_\ell} \pi(i) = 1\ .$$

- *If State j is recurrent and i is not in its irreducible set,*

$$\lim_{n\to\infty} \Pr\{X_n = j | X_0 = i\} = F(i,j)\pi(j)\ ,$$

  where $\boldsymbol{\pi}$ is determined as above.
- *If State j is recurrent and X_0 is in the same irreducible set as j,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} I(X_m, j) = \pi(j)\ ,$$

  where $\mathbf{I}$ is the identity matrix.
- *If State j is recurrent,*

$$E[T^j | X_0 = j] = \frac{1}{\pi(j)}\ .$$

The intuitive idea of the second to last item in the above property is obtained by considering the role that the identity matrix plays in the left-hand-side of the equation. As mentioned previously, the identity matrix acts as a counter so that the summation on the left-hand-side of the equation is the total number of visits that the chain makes to state j. Thus, the equality indicates that the fraction of time spent in State j is equal to the steady-state probability of being in State j. This property is called the *Ergodic Property*. The last property indicates that the reciprocal of the long-run probabilities equals the expected number of steps to return to the state. Intuitively, this is as one would expect it, since the higher the probability, the quicker the return.

Example 5.11. We consider again the Markov chain of Example 5.10. In order to determine its steady-state behavior, the first step is to rearrange the matrix as in Eq. (5.8). The irreducible sets were identified in Example 5.10 and on that basis we order the state space as b,g,c,d,f,a,e. The Markov matrix thus becomes

$$\mathbf{P} = \begin{array}{c} b \\ g \\ \hline c \\ d \\ f \\ \hline a \\ e \end{array} \left[\begin{array}{cc|ccc|cc} 0.5 & 0.5 & & & & & \\ 0.8 & 0.2 & & & & & \\ \hline & & 0.4 & 0.6 & 0 & & \\ & & 0.3 & 0.2 & 0.5 & & \\ & & 1 & 0 & 0 & & \\ \hline 0.1 & 0 & 0.2 & 0.2 & 0.1 & 0.3 & 0.1 \\ 0 & 0 & 0.2 & 0.3 & 0.1 & 0 & 0.4 \end{array}\right] .$$

(Blank blocks in a matrix are always interpreted to be zeroes.) The steady-state probabilities of being in States b or g are found by solving

$$\begin{aligned} 0.5\pi_b + 0.8\pi_g &= \pi_b \\ 0.5\pi_b + 0.2\pi_g &= \pi_g \\ \pi_b + \quad \pi_g &= 1 . \end{aligned}$$

Thus, after deleting one of the first two equations, we solve the two-by-two system and obtain $\pi_b = 8/13$ and $\pi_g = 5/13$. Of course if the chain starts in State c,d, or f the long-run probability of being in State b or g is zero, because it is impossible to reach the irreducible set $\{b,g\}$ from any state in the irreducible set $\{c,d,f\}$. The steady-state probabilities of being in State c,d, or f starting in that set is given by the solution to

$$\begin{aligned} 0.4\pi_c + 0.3\pi_d + \pi_f &= \pi_c \\ 0.6\pi_c + 0.2\pi_d \qquad &= \pi_d \\ 0.5\pi_d \qquad &= \pi_f \\ \pi_c + \quad \pi_d + \pi_f &= 1 \end{aligned}$$

which yields $\pi_c = 8/17$, $\pi_d = 6/17$, and $\pi_f = 3/17$. (It might be noticed that the reasonable person would delete the first equation before solving the above system.) As a final point, assume the chain is in State c, and we wish to know the expected number of steps until the chain returns to State c. Since $\pi_c = 8/17$, it follows that the expected number of steps until the first return is $17/8$. □

Example 5.12. The market analysis discussed in Example 5.5 gave switching probabilities for body styles and associated profits with each body style. The long behavior for a customer can be estimated by calculating the long-run probabilities. The following system of equations

$$\begin{aligned} 0.36\pi_s + 0.60\pi_w \quad &= \pi_w \\ 0.09\pi_s \quad + 0.25\pi_c &= \pi_c \\ \pi_s + \quad \pi_w + \quad \pi_c &= 1 \end{aligned}$$

is solved to obtain $\pi = (0.495, 0.446, 0.059)$. Therefore, the long-run expected profit per customer trade-in is

$$0.495 \times 1200 + 0.446 \times 1500 + 0.059 \times 2500 = \$1,410.5 \,.$$

□

- *Suggestion: Do Problems 5.10d, 5.11g–i, and 5.15.*

5.5 Computations

The determination of **R** (Eq. 5.5) is straightforward for recurrent states. If States i and j are in the same irreducible set, then $R(i,j) = \infty$. If i is recurrent and j is either transient or in a different irreducible set than i, then $R(i,j) = F(i,j) = 0$. For i transient and j recurrent, then $R(i,j) = \infty$ if $F(i,j) > 0$ and $R(i,j) = 0$ if $F(i,j) = 0$. In the case where i and j are transient, we have to do a little work as is given in the following property.

Property 5.11. *Let $X = \{X_n; n = 0,1,\cdots\}$ be a Markov chain and let A denote the (finite) set of all transient states. Let $\mathbf{Q}$ be the matrix of transition probabilities restricted to the set A, then, for $i, j \in A$*

$$R(i,j) = (\mathbf{I} - \mathbf{Q})^{-1}(i,j) \,.$$

Continuing with Example 5.11, we first note that $\mathbf{Q}$ is

$$\mathbf{Q} = \begin{bmatrix} 0.3 & 0.1 \\ 0 & 0.4 \end{bmatrix}$$

and evaluating $(\mathbf{I} - \mathbf{Q})^{-1}$ we have

$$(\mathbf{I} - \mathbf{Q})^{-1} = \begin{bmatrix} 10/7 & 10/42 \\ 0 & 10/6 \end{bmatrix}$$

which yields

$$
\mathbf{R} = \left[\begin{array}{cc|ccc|cc}
\infty & \infty & 0 & 0 & 0 & 0 & 0 \\
\infty & \infty & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & \infty & \infty & \infty & 0 & 0 \\
0 & 0 & \infty & \infty & \infty & 0 & 0 \\
0 & 0 & \infty & \infty & \infty & 0 & 0 \\
\hline
\infty & \infty & \infty & \infty & \infty & 10/7 & 10/42 \\
0 & 0 & \infty & \infty & \infty & 0 & 10/6
\end{array}\right] .
$$

The calculations for the matrix $\mathbf{F}$ are slightly more complicated than for $\mathbf{R}$. The matrix $\mathbf{P}$ must be rewritten so that each irreducible, recurrent set is treated as a single "super" state. Once the Markov chain gets into an irreducible, recurrent set, it will remain in that set forever and all states within the set will be visited infinitely often. Therefore, in determining the probability of reaching a recurrent state from a transient state, it is only necessary to find the probability of reaching the appropriate irreducible set. The transition matrix in which each irreducible set is treated as a single state is denoted by $\hat{\mathbf{P}}$. The matrix $\hat{\mathbf{P}}$ has the form

$$
\hat{\mathbf{P}} = \begin{bmatrix}
1 & & & & \\
 & 1 & & & \\
 & & 1 & & \\
 & & & \ddots & \\
\mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 & \cdots & \mathbf{Q}
\end{bmatrix} , \tag{5.9}
$$

where $\mathbf{b}_\ell$ is a vector giving the one-step probability of going from transient state i to irreducible set ℓ, that is,

$$
b_\ell(i) = \sum_{j \in C_\ell} P(i,j)
$$

for i transient and C_ℓ denoting the ℓ^{th} irreducible set.

Property 5.12. *Let $\{X_n; n = 0, 1, \cdots\}$ be a Markov chain with a finite state space ordered so that its Markov matrix can be reduced to the form in Eq. (5.9). Then for a transient State i and a recurrent State j, we have*

$$
F(i,j) = ((\mathbf{I} - \mathbf{Q})^{-1}\mathbf{b}_\ell)(i)
$$

for each State j in the ℓ^{th} irreducible set.

We again return to Example 5.11 to illustrate the calculations for $\mathbf{F}$. We will also take advantage of Property 5.5 to determine the values in the lower right-hand portion $\mathbf{F}$. Note that for i and j both transient, Property 5.5 can be rewritten as

$$
F(i,j) = \begin{cases} 1 - \frac{1}{R(j,j)} & \text{for } i = j, \\[2ex] \frac{R(i,j)}{R(j,j)} & \text{for } i \neq j\,. \end{cases} \tag{5.10}
$$

Using Property 5.12, the following holds:

$$\hat{\mathbf{P}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0.1 & 0.5 & 0.3 & 0.1 \\ 0 & 0.6 & 0 & 0.4 \end{bmatrix},$$

$$(\mathbf{I}-\mathbf{Q})^{-1}\mathbf{b}_1 = \begin{bmatrix} 1/7 \\ \\ 0 \end{bmatrix},$$

$$(\mathbf{I}-\mathbf{Q})^{-1}\mathbf{b}_2 = \begin{bmatrix} 6/7 \\ \\ 1 \end{bmatrix};$$

thus using Property 5.12 and Eq. (5.10), we have

$$\mathbf{F} = \left[\begin{array}{cc|ccc|cc} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ \hline 1/7 & 1/7 & 6/7 & 6/7 & 6/7 & 3/10 & 1/7 \\ 0 & 0 & 1 & 1 & 1 & 0 & 4/10 \end{array}\right].$$

We can now use Property 5.10 and the previously computed steady-state probabilities to finish the limiting probability matrix for this example; namely,

$$\lim_{n\to\infty} \mathbf{P}^n = \left[\begin{array}{cc|ccc|cc} 8/13 & 5/13 & 0 & 0 & 0 & 0 & 0 \\ 8/13 & 5/13 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 8/17 & 6/17 & 3/17 & 0 & 0 \\ 0 & 0 & 8/17 & 6/17 & 3/17 & 0 & 0 \\ 0 & 0 & 8/17 & 6/17 & 3/17 & 0 & 0 \\ \hline 8/91 & 5/91 & 48/119 & 36/119 & 18/119 & 0 & 0 \\ 0 & 0 & 8/17 & 6/17 & 3/17 & 0 & 0 \end{array}\right].$$

The goal of many modeling projects is the determination of revenues or costs. The ergodic property (the last item in Property 5.10) of an irreducible recurrent Markov chain gives an easy formula for the long run average return for the process. In Example 5.1, assume that each day the tractor is running, a profit of \$100 is realized; however, each day it is in the repair shop, a cost of \$25 is incurred. The Markov matrix of Example 5.1 yields steady-state results of $\pi_0 = 1/11$ and $\pi_1 = 10/11$; thus, the daily average return in the long run is

$$-25 \times \frac{1}{11} + 100 \times \frac{10}{11} = 88.64 \,.$$

This intuitive result is given in the following property.

Property 5.13. *Let $X = \{X_n; n = 0, 1, \cdots\}$ be an irreducible Markov chain with finite state space E and with steady-state probabilities given by the vector $\boldsymbol{\pi}$. Let the vector* **f** *be a profit function (i.e., $f(i)$ is the profit received for each visit to State i). Then (with probability one) the long-run average profit per unit of time is*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \sum_{j \in E} \pi(j) f(j) \,.$$

Example 5.13. We return to the simplified manufacturing system of Example 5.8. Refer again to the diagram in Fig. 5.7 and the Markov matrix of Eq. 5.6. The cost structure for the process is as follows: The cost of the raw material going into Step 1 is \$150; each time a part is processed through Step 1 a cost of \$200 is incurred; and every time a part is processed through Step 2 a cost of \$300 is incurred. (Thus if a part is sold that was reworked once in Step 1 but was not reworked in Step 2, that part would have \$850 of costs associated with it.) Because the raw material is toxic, there is also a disposal cost of \$50 per part sent to scrap.

Each day, we start with enough raw material to make 100 parts so that at the end of each day, 100 parts are finished: some good, some scraped. In order to establish a reasonable strategy for setting the price of the parts to be sold, we first must determine the cost that should be attributed to the parts. In order to answer the relevant questions, we will first need the **F** and **R** matrices, which are

$$\mathbf{F} = \begin{array}{c} 1 \\ 2 \\ s \\ g \end{array} \begin{bmatrix} 0.239 & 0.875 & 0.183 & 0.817 \\ 0.056 & 0.144 & 0.066 & 0.934 \\ 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$\mathbf{R} = \begin{array}{c} 1 \\ 2 \\ s \\ g \end{array} \begin{bmatrix} 1.31 & 1.02 & \infty & \infty \\ 0.07 & 1.17 & \infty & \infty \\ 0.0 & 0.0 & \infty & 0.0 \\ 0.0 & 0.0 & 0.0 & \infty \end{bmatrix} .$$

An underlying assumption for this model is that each part that begins the manufacturing process is an independent entity, so that the actual number of finished "good" parts is a random variable following a binomial distribution. As you recall, the binomial distribution needs two parameters, the number of trials and the probability of success. The number of trials is given in the problem statement as 100; the probability of a success is the probability that a part which starts in State 1 ends in State g. Therefore, the expected number of good parts at the end of the day is

$$100 \times F(1,g) = 81.7 .$$

The cost per part started is given by

$$150 + 200 \times R(1,1) + 300 \times R(1,2) + 50 \times F(1,s) = 727.15 .$$

Therefore, the cost which should be associated to each part sold is

$$(100 \times 727.15)/81.7 = \$890/\text{part sold} .$$

A rush order for 100 parts from a very important customer has just been received, and there are no parts in inventory. Therefore, we wish to start tomorrow's production with enough raw material to be 95% confident that there will be at least 100 good parts at the end of the day. How many parts should we plan on starting tomorrow morning? Let the random variable N_n denote the number of finished parts that are good given that the day started with enough raw material for n parts. From the above discussion, the random variable N_n has a binomial distribution where n is the number of trials and $F(1,g)$ is the probability of success. Therefore, the question of interest is to find n such that $\Pr\{N_n \geq 100\} = 0.95$. Hopefully, you also recall that the binomial distribution can be approximated by the normal distribution; therefore, define X to be a normally distributed random variable with mean $nF(1,g)$ and variance $nF(1,g)F(1,s)$. We now have the following equation:

$$\begin{aligned}\Pr\{N_n \geq 100\} &\approx \Pr\{X > 99.5\} \\ &= \Pr\{Z > (99.5 - 0.817n)/\sqrt{0.1495n}\} = 0.95\end{aligned}$$

where Z is normally distributed with mean zero and variance one. Thus, using either standard statistical tables or the Excel function `=NORMSINV(0.05)`, we have that

$$\frac{99.5 - 0.817n}{\sqrt{0.1495n}} = -1.645 .$$

The above equation is solved for n. Since it becomes a quadratic equation, there are two roots: $n_1 = 113.5$ and $n_2 = 130.7$. We must take the second root (why?), and we round up; thus, the day must start with enough raw material for 131 parts. □

Example 5.14. A missile is launched and, as it is tracked, a sequence of course correction signals are sent to it. Suppose that the system has four states that are labeled as follows.

State 0: on-course, no further correction necessary
State 1: minor deviation
State 2: major deviation
State 3: abort, off-course so badly a self-destruct signal is sent

Let X_n represent the state of the system after the n^{th} course correction and assume that the behavior of X can be modeled by a Markov chain with the following probability transition matrix:

$$\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 \\ 0.5 & 0.25 & 0.25 & 0.0 \\ 0.0 & 0.5 & 0.25 & 0.25 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix} .$$

As always, the first step is to classify the states. Therefore observe that states 0 and 3 are absorbing, and states 1 and 2 are transient. After a little work, you should be able to obtain the following matrices:

$$\mathbf{F} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 \\ 0.857 & 0.417 & 0.333 & 0.143 \\ 0.571 & 0.667 & 0.417 & 0.429 \\ 0.0 & 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$\mathbf{R} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{bmatrix} \infty & 0.0 & 0.0 & 0.0 \\ \infty & 1.714 & 0.571 & \infty \\ \infty & 1.143 & 1.714 & \infty \\ 0.0 & 0.0 & 0.0 & \infty \end{bmatrix} .$$

Suppose that upon launch, the missile starts in State 2. The probability that it will eventually get on-course is 57.1% (namely, F(2,0)); whereas, the probability that it will eventually have to be destroyed is 42.9%. When a missile is launched, 50,000 pounds of fuel are used. Every time a minor correction is made, 1,000 pounds of fuel are used; and every time a major correction is made, 5,000 pounds of fuel are used. Assume that the missile started in State 2, we wish to determine the expected fuel usage for the mission. This calculation is

$$50000 + 1000 \times R(2,1) + 5000 \times R(2,2) = 59713 .$$

□

The decision maker sometimes uses a discounted cost criterion instead of average costs over an infinite planning horizon. To determine the expected total discounted return for a Markov chain, we have the following.

Property 5.14. *Let $X = \{X_n; n = 0, 1, \cdots\}$ be a Markov chain with Markov matrix $\mathbf{P}$. Let $\mathbf{f}$ be a return function and let α $(0 < \alpha < 1)$ be a discount factor. Then the expected total discounted cost is given by*

$$E\left[\sum_{n=0}^{\infty} \alpha^n f(X_n) | X_0 = i\right] = ((\mathbf{I} - \alpha\mathbf{P})^{-1}\mathbf{f})(i) .$$

We finish Example 5.11 by considering a profit function associated with it. Let the state space for the example be $E = \{1,2,3,4,5,6,7\}$; that is, State 1 is State b, State 2 is State g, etc. Let $\mathbf{f}$ be the profit function with $f = (500, 450, 400, 350, 300, 350, 200)$; that is, each visit to State 1 produces \$500, each visit to State 2 produces \$450, etc. If the chain starts in States 1 or 2, then the long-run average profit per period is

$$500 \times \frac{8}{13} + 450 \times \frac{5}{13} = \$480.77 \, .$$

If the chain starts in States 3, 4, or 5, then the long-run average profit per period is

$$400 \times \frac{8}{17} + 350 \times \frac{6}{17} + 300 \times \frac{3}{17} = \$364.71 \, .$$

If the chain starts in State 6, then the expected value of the long-run average would be a weighted average of the two ergodic results, namely,

$$480.77 \times \frac{1}{7} + 364.71 \times \frac{6}{7} = \$381.29 \, ,$$

and starting in State 7 gives the same long-run result as starting in States 3, 4, or 5.

Now assume that a discount factor should be used. In other words, our criterion is total discounted profit instead of long-run average profit per period. Let the monetary rate of return be 20% per period which gives $\alpha = 1/1.20 = 5/6$. (In other words, one dollar one period in the future is equivalent to 5/6 dollar, or 83.33 cents, now.) Then

$$(\mathbf{I} - \alpha \mathbf{P})^{-1} = \left[\begin{array}{cc|ccc|cc} 4.0 & 2.0 & 0 & 0 & 0 & 0 & 0 \\ 3.2 & 2.8 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 3.24 & 1.95 & 0.81 & 0 & 0 \\ 0 & 0 & 2.32 & 2.59 & 1.08 & 0 & 0 \\ 0 & 0 & 2.70 & 1.62 & 1.68 & 0 & 0 \\ \hline 0.44 & 0.22 & 1.76 & 1.37 & 0.70 & 1.33 & 0.17 \\ 0 & 0 & 2.02 & 1.66 & 0.82 & 0 & 1.50 \end{array} \right] .$$

Denote the total discounted return when the chain starts in state i by $h(i)$. Then, multiplying the above matrix by the profit function $\mathbf{f}$ yields

$$\mathbf{h}^T = (2900, 2860, 2221.5, 2158.5, 2151, 2079, 1935) \, .$$

Thus, if the chain started in State 1, the present worth of all future profits would be \$2,900. If the initial state is unknown and a distribution of probabilities for the starting state is given by $\boldsymbol{\mu}$, the total discounted profit is given by $\boldsymbol{\mu}\mathbf{h}$.

- *Suggestion: Do Problems 5.13d–f, 5.14e–g, and finish 5.1–5.9.*

Appendix

We close the chapter with a brief look at the use of Excel for analyzing Markov chains. We will consider both the matrix operations involved in the computations of Section 5.5 and the use of random numbers for simulating Markov chains.

Matrix Operations in Excel. The application of Property 5.11 to the Markov chain of Example 5.11 requires the determination of $(\mathbf{I}-\mathbf{Q})^{-1}$ as shown on p. 162. There is more than one way to calculate this in Excel, but they all involve the use of the function MINVERSE(*array*). In Excel, whenever the output of a function is an array it is important to follow Excel's special instructions for array results: (1) highlight the cells in which you want the array to appear, (2) type in the equal sign (=) followed by the formula, and (3) hold down the <shift> and <ctrl> keys while hitting the <enter> key. To demonstrate this, type the following on a spreadsheet:

	A	B	C	D	E
1	Q matrix			Identity	
2	0.3	0.1		1	0
3	0.0	0.4		0	1

(Notice that as a matter of personal preference, parameters are usually labeled and blank columns or rows are placed between separate quantities.) Now to obtain the inverse of $(\mathbf{I}-\mathbf{Q})$, first, highlight the cells A5:B6; second, type

=MINVERSE(D2:E3 - A2:B3)

(note that although the range A5:B6 was highlighted, the above formula should occur in Cell A5); finally, while holding down the <shift> and <ctrl> keys hit the <enter> key. After hitting the <enter> key, the $(\mathbf{I}-\mathbf{Q})^{-1}$ matrix should appear in the A5:B6 range. To try an alternate method, highlight the cells A8:B9, type

=MINVERSE({1,0;0,1} - A2:B3), and

while holding down the <shift> and <ctrl> keys hit the <enter> key. The obvious advantage of the second method is that the identity matrix need not be explicitly entered on the spreadsheet.

Steady-state probabilities for an irreducible, recurrent Markov chain can also be easily obtained through Excel, but first we re-write the equation $\boldsymbol{\pi}\mathbf{P}=\boldsymbol{\pi}$ as $\boldsymbol{\pi}(\mathbf{I}-\mathbf{P})=\mathbf{0}$. Recall from Property 5.9, that the norming equation must be used and one of the equations from the system $\boldsymbol{\pi}(\mathbf{I}-\mathbf{P})=\mathbf{0}$ must be dropped, which is equivalent to deleting one of the *columns* of $(\mathbf{I}-\mathbf{P})$ and replacing it with a column of all ones. As we do this in Excel, it is the last column that will be dropped.

Example 5.15. Consider a four-state Markov chain with Markov matrix defined by Eq. (5.7). Since it has a finite state space and all states communicate with all other states, it is an irreducible, recurrent Markov chain so that steady-state probabilities can be computed using Property 5.9. Type the following in a spreadsheet:

	A	B	C	D
1		P matrix		
2	0.5	0.5	0	0
3	0.2	0.4	0.4	0
4	0	0.1	0.8	0.1
5	0	0	0.3	0.7
6				
7		Modified I-P matrix		

Next, we need the matrix $(\mathbf{I}-\mathbf{P})$ with the last column replaced by all ones to appear in Cells A8:D11. To accomplish this, highlight Cells A8:C11, type

```
={1,0,0;0,1,0;0,0,1;0,0,0}-A2:C5
```

(again note that although a block of 12 cells were highlighted, the typing occurs in Cell A8), and while holding down the `<shift>` and `<ctrl>` keys hit the `<enter>` key. Finally, place ones in Cells D8:D11.

With the modified $(\mathbf{I}-\mathbf{P})$ matrix, we simply pre-multiply the right-hand side vector by its inverse, where the right-hand side is a vector of all zeros except the last element which equals one (from the norming equation). Center the words `"Steady state probabilities"` across Cells A13:D13 in order to label the quantities to be computed, highlight Cells A14:D14, type

```
=MMULT({0,0,0,1},MINVERSE(A8:D11))
```

(although a block of 4 cells were highlighted, the typing occurs in Cell A14), and while holding down the `<shift>` and `<ctrl>` keys hit the `<enter>` key. Thus, the steady-state probabilities appear in highlighted cells. □

Simulation of a Markov Chain. We close this appendix with an Excel simulation of a Markov chain. The key to simulating a Markov chain is to use the techniques of Sect. 2.3.1 together with conditional probabilities given in the Markov matrix. The Excel implementation of this was illustrated in Example 2.12.

Example 5.16. Consider a Markov chain with state space $\{a,b,c\}$ and with the following transition matrix:

$$\mathbf{P} = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.4 & 0.6 & 0.0 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}.$$

We know that the expected number of steps taken until the chain is absorbed into State c given that it starts in State a is $R(a,a)+R(a,b) = 12.5$ (Eq. 5.5). However, let us assume that as a new student of Markov chains, our confidence level is low in asserting that the mean absorption time starting from State a is 12.5 (or equivalently, $E[T^c|X_0 = a] = 12.5$). One of the uses of simulation is as a confidence builder; therefore, we wish to simulate this process to help give some evidence that our calculations are correct. In other words, our goal is to simulate this Markov chain starting from State a and recording the number of steps taken until reaching State c.

Prepare a spreadsheet so that the first two rows are as follows:

	A	B	C	D
1	`Time Step`	`Current State`	`Random Number`	`Number of Steps`
2	`0`	`a`	`=RAND()`	`=MATCH("c",B3:B200, 0)`

For Cell A3, type `=A2+1`; for Cell B3, type

```
=IF(B2="a",IF(C2<0.8,"a",IF(C2<0.9,"b","c")),
      IF(B2="b",IF(C2<0.4,"a","b"),"c"))
```

and for Cell C3, type `=RAND()`. Now copy Cells A3:C3 down through row 200.

Notice that each time the F9 key is pressed, a new value appears in Cell D2. These values represent independent realizations of the random variable T^c given that $X_0 = a$. Obviously, to obtain an estimate for $E[T^c|X_0 = a]$, an average of several realizations are necessary. □

The above example can be modified slightly so that one column produces a random sample containing many realizations of $E[T^c|X_0 = a]$. Change Cell A3 to read `=IF(B2="c",0,A2+1)` (note that Cell A3 makes reference to B2, not B3), change the last entry in the `IF` statement of Cell B3 to be `"a"` instead of `"c"`, change Cell D3 to be `=IF(B3="c", A3,"--")`, and then copy Cells A3:D3 down through row 30000. The numerical values contained in Column D represent various realizations of the random variable T^c so we now change the formula in D2 to be `=AVERAGE(D3:D30000)` and thus the sample mean becomes the estimator for $E[T^c|X_0 = a]$.

Problems

The exercises listed below are to help in your understanding of Markov chains and to indicate some of the potential uses for Markov chain analysis. All numbers are fictitious.

5.1. Joe and Pete each have two cents in their pockets. They have decided to match pennies; that is, they will each take one of their own pennies and flip them. If the pennies match (two heads or two tails), Joe gets Pete's penny; if the pennies do not match, Pete gets Joe's penny. They will keep repeating the game until one of them has four cents, and the other one is broke. Although they do not realize it, all four pennies are biased. The probability of tossing a head is 0.6, and the probability of a tail is 0.4. Let X be a Markov chain where X_n denotes the amount that Joe has after the n^{th} play of the game.
(a) Give the Markov matrix for X.
(b) What is the probability that Joe will have four pennies after the second toss?
(c) What is the probability that Pete will be broke after three tosses?
(d) What is the probability that the game will be over before the third toss?
(e) What is the expected amount of money Pete will have after two tosses?

(f) What is the probability that Pete will end up broke?
(g) What is the expected number of tosses until the game is over?

5.2. At the start of each week, the condition of a machine is determined by measuring the amount of electrical current it uses. According to its amperage reading, the machine is categorized as being in one of the following four states: low, medium, high, failed. A machine in the low state has a probability of 0.05, 0.03, and 0.02 of being in the medium, high, or failed state, respectively, at the start of the next week. A machine in the medium state has a probability of 0.09 and 0.06 of being in the high or failed state, respectively, at the start of the next week (it cannot, by itself, go to the low state). And, a machine in the high state has a probability of 0.1 of being in the failed state at the start of the next week (it cannot, by itself, go to the low or medium state). If a machine is in the failed state at the start of a week, repair is immediately begun on the machine so that it will (with probability 1) be in the low state at the start of the following week. Let X be a Markov chain where X_n is the state of the machine at the start of week n.
(a) Give the Markov matrix for X.
(b) A new machine always starts in the low state. What is the probability that the machine is in the failed state three weeks after it is new?
(c) What is the probability that a machine has at least one failure three weeks after it is new?
(d) On the average, how many weeks per year is the machine working?
(e) Each week that the machine is in the low state, a profit of \$1,000 is realized; each week that the machine is in the medium state, a profit of \$500 is realized; each week that the machine is in the high state, a profit of \$400 is realized; and the week in which a failure is fixed, a cost of \$700 is incurred. What is the long-run average profit per week realized by the machine?
(f) A suggestion has been made to change the maintenance policy for the machine. If at the start of a week the machine is in the high state, the machine will be taken out of service and repaired so that at the start of the next week it will again be in the low state. When a repair is made due to the machine being in the high state instead of a failed state, a cost of \$600 is incurred. Is this new policy worthwhile?

5.3. We are interested in the movement of patients within a hospital. For purposes of our analysis, we shall consider the hospital to have three different types of rooms: general care, special care, and intensive care. Based on past data, 60% of arriving patients are initially admitted into the general care category, 30% in the special care category, and 10% in intensive care. A "general care" patient has a 55% chance of being released healthy the following day, a 30% of remaining in the general care room, and a 15% of being moved to the special care facility. A "special care" patient has a 10% chance of being released the following day, a 20% chance of being moved to general care, a 10% chance of being upgraded to intensive care, and a 5% chance of dying during the day. An "intensive care" patient is never released from the hospital directly from the intensive care unit (ICU), but is always moved to another facility first. The probabilities that the patient is moved to general care, special care, or remains in intensive care are 5%, 30%, or 55%, respectively. Let X

be a Markov chain where X_0 is the type of room that an admitted patient initially uses, and X_n is the room category of that patient at the end of day n.
(a) Give the Markov matrix for X.
(b) What is the probability that a patient admitted into the intensive care room eventually leaves the hospital healthy?
(c) What is the expected number of days that a patient, admitted into intensive care, will spend in the ICU?
(d) What is the expected length of stay for a patient admitted into the hospital as a general care patient?
(e) During a typical day, 100 patients are admitted into the hospital. What is the average number of patients in the ICU?

5.4. Consider again the manufacturing process of Example 5.8. New production plans call for an expected production level of 2000 good parts per month. (In other words, enough raw material must be used so that the expected number of good parts produced each month is 2000.) For a capital investment and an increase in operating costs, all rework and scrap can be eliminated. The sum of the capital investment and operating cost increase is equivalent to an annual cost of $5 million. Is it worthwhile to increase the annual cost by $5 million in order to eliminate the scrap and rework?

5.5. Assume the description of the manufacturing process of Example 5.8 is for the process at Branch A of the manufacturing company. Branch B of the same company has been closed. Branch B had an identical process and they had 1000 items in stock that had been through Step 1 of the process when they were shut down. Because the process was identical, these items can be fed into Branch A's process at the start of Step 2. (However, since the process was identical there is still a 5% chance that after finishing Step 2 the item will have to be reworked at Step 1, a 10% chance the item will have to be reworked at Step 2, and a 5% chance that the item will have to be scrapped.) Branch A purchases these (partially completed) items for a total of $300,000, and they will start processing at Step 2 of Branch A's system. After Branch A finishes processing this batch of items, they must determine the cost of these items so they will know how much to charge customers in order to recover the cost. (They may want to give a discount.) Your task is to determine the cost that would be attributed to each item shipped.

5.6. The manufacture of a certain type of electronic board consists of four steps: tinning, forming, insertion, and solder. After the forming step, 5% of the parts must be retinned; after the insertion step, 20% of the parts are bad and must be scrapped; and after the solder step, 30% of the parts must be returned to insertion and 10% must be scrapped. (We assume that when a part is returned to a processing step, it is treated like any other part entering that step.) Figure 5.12 gives a schematic showing the flow of a job through the manufacturing steps.
(a) Model this process using Markov chains and give its Markov matrix.
(b) If a batch of 100 boards begins this manufacturing process, what is the expected number that will end up scraped?
(c) How many boards should we start with if the goal is to have the expected number

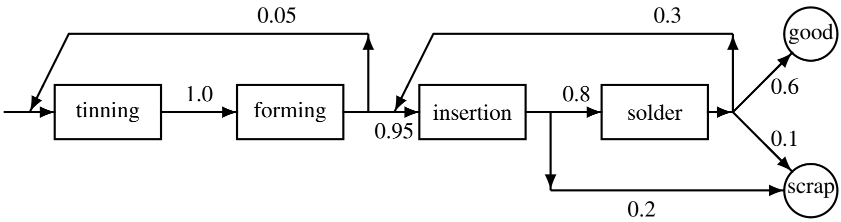


Fig. 5.12 *Manufacturing structure for board processing steps*

of boards that finish in the good category equal to 100?
(d) How many boards should we start with if we want to be 90% sure that we end up with a batch of 100 boards? (Hint: the final status of each board can be considered a Bernoulli random variable, the sum of independent Bernoullis is binomial, and the binomial can be approximated by a normal.)
(e) Each time a board goes through a processing step, direct labor and material costs are \$10 for tinning, \$15 for forming, \$25 for insertion, and \$20 for solder. The raw material costs \$8, and a scrapped board returns \$2. The average overhead rate is \$1,000,000 per year, which includes values for capital recovery. The average processing rate is 5,000 board starts per week. We would like to set a price per board so that expected revenues are 25% higher than expected costs. At what value would you set the price per board?

5.7. The government has done a study of the flow of money among three types of financial institutions: banks, savings associations, and credit unions. We shall assume that the Markov assumptions are valid and model the monthly movement of money among the various institutions. Some recent data are:

June amounts	Units in billions of dollars May				June totals
	bank	savings	credit	other	
bank	10.0	4.5	1.5	3.6	19.6
savings	4.0	6.0	0.5	1.2	11.7
credit	2.0	3.0	6.0	2.4	13.4
other	4.0	1.5	2.0	4.8	12.3
May totals	20	15	10	12	57

For example, of the 20 billion dollars in banks during the month of May, half of it remained in banks for the month of June, and 2 out of the 20 billion left banks to be invested in credit unions during June.
(a) Use the above data to estimate a Markov matrix. What would be the problems with such a model? (That is, be critical of the Markov assumption of the model.)
(b) In the long-run, how much money would you expect to be in credit unions during any given month?

(c) How much money would you expect to be in banks during August of the same year that the above data were collected?

5.8. Within a certain market area, there are two brands of soap that most people use: "super soap" and "cheap soap", with the current market split evenly between the two brands. A company is considering introducing a third brand called "extra clean soap", and they have done some initial studies of market conditions. Their estimates of weekly shopping patterns are as follows: If a customer buys super soap this week, there is a 75% chance that next week the super soap will be used again, a 10% chance that extra clean will be used and a 15% chance that the cheap soap will be used. If a customer buys the extra clean this week, there is a fifty-fifty chance the customer will switch, and if a switch is made it will always be to super soap. If a customer buys cheap soap this week, it is equally likely that next week the customer will buy any of the three brands.
(a) Assuming that the Markov assumptions are good, use the above data to estimate a Markov matrix. What would be the problems with such a model?
(b) What is the long-run market share for the new soap?
(c) What will be the market share of the new soap two weeks after it is introduced?
(d) The market consists of approximately one million customers each week. Each purchase of super soap yields a profit of 15 cents; a purchase of cheap soap yields a profit of 10 cents; and a purchase of extra clean will yield a profit of 25 cents. Assume that the market was at steady-state with the even split between the two products. The initial advertising campaign to introduce the new brand was $100,000. How many weeks will it be until the $100,000 is recovered from the added revenue of the new product?
(e) The company feels that with these three brands, an advertising campaign of $30,000 per week will increase the weekly total market by a quarter of a million customers? Is the campaign worthwhile? (Use a long-term average criterion.)

5.9. A small company sells high quality laser printers and they use a simple periodic inventory ordering policy. If there are two or fewer printers in inventory at the end of the day on Friday, the company will order enough printers so that there will be five printers in stock at the start of Monday. (It only takes the weekend for printers to be delivered from the wholesaler.) If there are more than two printers at the end of the week, no order is placed. Weekly demand data has been analyzed yielding the probability mass function for weekly demand as $\Pr\{D=0\}=0.05$, $\Pr\{D=1\}=0.1$, $\Pr\{D=2\}=0.2$, $\Pr\{D=3\}=0.4$, $\Pr\{D=4\}=0.1$, $\Pr\{D=5\}=0.1$, and $\Pr\{D=6\}=0.05$. Let X be a Markov chain where X_n is the inventory at the end of week n. (Note: backorders are not accepted.)
(a) Give the Markov matrix for X.
(b) If at the end of week 1 there were five items in inventory, what is the probability that there will be five items in inventory at the end of week 2?
(c) If at the end of week 1 there were five items in inventory, what is the probability that there will be five items in inventory at the end of week 3?
(d) What is the expected number of times each year that an order is placed?
(e) Each printer sells for $1,800. Each time an order is placed, it costs $500 plus

\$1,000 times the number of items ordered. At the end of each week, each unsold printer costs \$25 (in terms of keeping them clean, money tied up, and other such inventory type expenses). Whenever a customer wants a printer not in stock, the company buys it retail and sends it by airfreight to the customer; thus the customer spends the \$1,800 sales price but it costs the company \$1,900. In order reduce the lost sales, the company is considering raising the reorder point to three, but still keeping the order up to quantity at five. Would you recommend the change in the reorder point?

5.10. Let X be a Markov chain with state space $\{a,b,c\}$ and transition probabilities given by

$$\mathbf{P} = \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 1.0 & 0.0 & 0.0 \\ 0.0 & 0.3 & 0.7 \end{bmatrix}.$$

(a) What is $\Pr\{X_2 = a | X_1 = b\}$?
(b) What is $\Pr\{X_2 = a | X_1 = b, X_0 = c\}$?
(c) What is $\Pr\{X_{35} = a | X_{33} = a\}$?
(d) What is $\Pr\{X_{200} = a | X_0 = b\}$? (Use steady-state probability to answer this.)

5.11. Let X be a Markov chain with state space $\{a,b,c\}$ and transition probabilities given by

$$\mathbf{P} = \begin{bmatrix} 0.3 & 0.7 & 0.0 \\ 0.0 & 0.6 & 0.4 \\ 0.4 & 0.1 & 0.5 \end{bmatrix}.$$

Let the initial probabilities be given by the vector $(0.1, 0.3, 0.6)$ with a profit function given by $(10, 20, 30)$. (That means, for example each visit to state a yields a profit of \$10.) Find the following:
(a) $\Pr\{X_2 = b | X_1 = c\}$.
(b) $\Pr\{X_3 = b | X_1 = c\}$.
(c) $\Pr\{X_3 = b | X_1 = c, X_0 = c\}$.
(d) $\Pr\{X_2 = b\}$.
(e) $\Pr\{X_1 = b, X_2 = c | X_0 = c\}$.
(f) $E[f(X_2) | X_1 = c]$.
(g) $\lim_{n\to\infty} \Pr\{X_n = a | X_0 = a\}$.
(h) $\lim_{n\to\infty} \Pr\{X_n = b | X_{10} = c\}$.
(i) $\lim_{n\to\infty} E[f(X_n) | X_0 = c]$.

5.12. Let X be a Markov chain with state space $\{a,b,c,d\}$ and transition probabilities given by

$$\mathbf{P} = \begin{bmatrix} 0.1 & 0.3 & 0.6 & 0.0 \\ 0.0 & 0.2 & 0.5 & 0.3 \\ 0.5 & 0.0 & 0.0 & 0.5 \\ 0.0 & 1.0 & 0.0 & 0.0 \end{bmatrix}.$$

Each time the chain is in state a, a profit of \$20 is made; each visit to state b yields a \$5 profit; each visit to state c yields \$15 profit; and each visit to state d costs \$10.

Find the following:
(a) $E[f(X_5)|X_3 = c, X_4 = d]$.
(b) $E[f(X_1)|X_0 = b]$.
(c) $E[f(X_1)^2|X_0 = b]$.
(d) $V[f(X_1)|X_0 = b]$.
(e) $V[f(X_1)]$ with an initial probability vector of $\boldsymbol{\mu} = (0.2, 0.4, 0.3, 0.1)$.

5.13. Consider the following Markov matrix representing a Markov chain with state space $\{a, b, c, d, e\}$.

$$\mathbf{P} = \begin{bmatrix} 0.3 & 0.0 & 0.0 & 0.7 & 0.0 \\ 0.0 & 1.0 & 0.0 & 0.0 & 0.0 \\ 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.5 & 0.5 & 0.0 \\ 0.0 & 0.2 & 0.4 & 0.0 & 0.4 \end{bmatrix}.$$

(a) Draw the state diagram.
(b) List the transient states.
(c) List the irreducible set(s).
(d) Let $F(i, j)$ denote the first passage probabilities of reaching (or returning to) State j given that $X_0 = i$. Calculate the $\mathbf{F}$ matrix.
(e) Let $R(i, j)$ denote the expected number of visits to State j given that $X_0 = i$. Calculate the $\mathbf{R}$ matrix.
(f) Calculate the $\lim_{n\to\infty} \mathbf{P}^n$ matrix.

5.14. Let X be a Markov chain with state space $\{a, b, c, d, e, f\}$ and transition probabilities given by

$$\mathbf{P} = \begin{bmatrix} 0.3 & 0.5 & 0.0 & 0.0 & 0.0 & 0.2 \\ 0.0 & 0.5 & 0.0 & 0.5 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.3 & 0.0 & 0.0 & 0.0 & 0.7 \\ 0.1 & 0.0 & 0.1 & 0.0 & 0.8 & 0.0 \\ 0.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}.$$

(a) Draw the state diagram.
(b) List the recurrent states.
(c) List the irreducible set(s).
(d) List the transient states.
(e) Calculate the $\mathbf{F}$ matrix.
(f) Calculate the $\mathbf{R}$ matrix.
(g) Calculate the $\lim_{n\to\infty} \mathbf{P}^n$ matrix.

5.15. Let X be a Markov chain with state space $\{a, b, c, d\}$ and transition probabilities given by

$$\mathbf{P} = \begin{bmatrix} 0.3 & 0.7 & 0.0 & 0.0 \\ 0.5 & 0.5 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 \\ 0.1 & 0.1 & 0.2 & 0.6 \end{bmatrix} .$$

(a) Find $\lim_{n\to\infty} \Pr\{X_n = a|X_0 = a\}$.
(b) Find $\lim_{n\to\infty} \Pr\{X_n = a|X_0 = b\}$.
(c) Find $\lim_{n\to\infty} \Pr\{X_n = a|X_0 = c\}$.
(d) Find $\lim_{n\to\infty} \Pr\{X_n = a|X_0 = d\}$.

5.16. Consider a Markov chain with state space $\{a,b,c\}$ and with the following transition matrix:

$$P = \begin{bmatrix} 0.35 & 0.27 & 0.38 \\ 0.82 & 0.00 & 0.18 \\ 1.00 & 0.00 & 0.00 \end{bmatrix} .$$

(a) Given that the Markov chain starts in State b, estimate through simulation the expected number of steps until the first return to State b.
(b) The expected return time to a state should equal the reciprocal of the long-run probability of being in that state. Estimate through simulation and analytically the steady-state probability and compare it to your answer to Part (a). Explain any differences.

5.17. Consider a Markov chain with state space $\{a,b,c\}$ and with the following transition matrix:

$$P = \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.1 & 0.2 & 0.7 \\ 0.8 & 0.0 & 0.2 \end{bmatrix} .$$

Each visit to State a results in a profit of \$5, each visit to State b results in a profit of \$10, and each visit to State c results in a profit of \$12. Write a computer program that will simulate the Markov chain so that an estimate of the expected profit per step can be made. Assume that the chain always starts in State a. The simulation should involve accumulating the profit from each step; then the estimate per simulation run is the cumulative profit divided by the number of steps in the run.
(a) Let the estimate be based on 10 replications, where each replication has 25 steps. The estimate is the average value over the 10 replicates. Record both the overall average and the range of the averages.
(b) Let the estimate be based on 10 replications, where each replication has 1000 steps. Compare the estimates and ranges for parts (a) and (b) and explain any differences.

References

1. Galton, F., and Watson, H.W. (1875). On the Probability of the Extinction of Families, *J. Anthropol. Soc. London*, **4**:138–144.

2. Miller, G.A. (1952). Finite Markov Processes in Psychology, *Psychometrika* **17**:149–167.
3. Parzen, E. (1962). *Stochastic Processes*, Holden-Day, Inc., Oakland, CA.