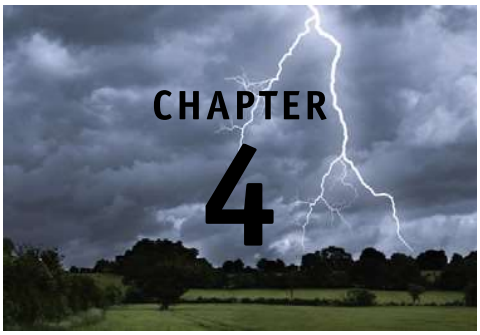
Markov Chains



4.1 Introduction

Consider a process that has a value in each time period. Let X_n denote its value in time period n, and suppose we want to make a probability model for the sequence of successive values $X_0, X_1, X_2 \ldots$. The simplest model would probably be to assume that the X_n are independent random variables, but often such an assumption is clearly unjustified. For instance, starting at some time suppose that X_n represents the price of one share of some security, such as Google, at the end of n additional trading days. Then it certainly seems unreasonable to suppose that the price at the end of day $n+1$ is independent of the prices on days $n, n-1, n-2$ and so on down to day 0. However, it might be reasonable to suppose that the price at the end of trading day $n+1$ depends on the previous end-of-day prices only through the price at the end of day n. That is, it might be reasonable to assume that the conditional distribution of X_{n+1} given all the past end-of-day prices $X_n, X_n - 1, \ldots, X_0$ depends on these past prices only through the price at the end of day n. Such an assumption defines a Markov chain, a type of stochastic process that will be studied in this chapter, and which we now formally define.

Let $\{X_n, n = 0, 1, 2, \ldots, \}$ be a stochastic process that takes on a finite or countable number of possible values. Unless otherwise mentioned, this set of possible values of the process will be denoted by the set of nonnegative integers $\{0, 1, 2, \ldots\}$. If $X_n = i$, then the process is said to be in state i at time n. We suppose that whenever the process is in state i, there is a fixed probability P_{ij} that it will next be in state j. That is, we suppose that

$$P\{X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1, X_0 = i_0\} = P_{ij} \tag{4.1}$$

Introduction to Probability Models, Eleventh Edition. http://dx.doi.org/10.1016/B978-0-12-407948-9.00004-9
© 2014 Elsevier Inc. All rights reserved.

for all states $i_0, i_1, \ldots, i_{n-1}, i, j$ and all $n \geqslant 0$. Such a stochastic process is known as a *Markov chain*. Equation (4.1) may be interpreted as stating that, for a Markov chain, the conditional distribution of any future state X_{n+1}, given the past states $X_0, X_1, \ldots, X_{n-1}$ and the present state X_n, is independent of the past states and depends only on the present state.

The value P_{ij} represents the probability that the process will, when in state i, next make a transition into state j. Since probabilities are nonnegative and since the process must make a transition into some state, we have

$$P_{ij} \geqslant 0, \qquad i, j \geqslant 0; \qquad \sum_{j=0}^{\infty} P_{ij} = 1, \qquad i = 0, 1, \ldots$$

kisi n kisi state me to jayega.. that's why sum is 1

Let $\mathbf{P}$ denote the matrix of one-step transition probabilities P_{ij}, so that

$$\mathbf{P} = \begin{Vmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ \vdots & \vdots & \vdots & \\ P_{i0} & P_{i1} & P_{i2} & \cdots \\ \vdots & \vdots & \vdots & \end{Vmatrix}$$

Example 4.1 (Forecasting the Weather) Suppose that the chance of rain tomorrow depends on previous weather conditions only through whether or not it is raining today and not on past weather conditions. Suppose also that if it rains today, then it will rain tomorrow with probability α; and if it does not rain today, then it will rain tomorrow with probability β.

If we say that the process is in state 0 when it rains and state 1 when it does not rain, then the preceding is a two-state Markov chain whose transition probabilities are given by

$$\mathbf{P} = \begin{Vmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{Vmatrix} \qquad \blacksquare$$

Example 4.2 (A Communications System) Consider a communications system that transmits the digits 0 and 1. Each digit transmitted must pass through several stages, at each of which there is a probability p that the digit entered will be unchanged when it leaves. Letting X_n denote the digit entering the nth stage, then $\{X_n, n = 0, 1, \ldots\}$ is a two-state Markov chain having a transition probability matrix

$$\mathbf{P} = \begin{Vmatrix} p & 1-p \\ 1-p & p \end{Vmatrix} \qquad \blacksquare$$

Example 4.3 On any given day Gary is either cheerful (C), so-so (S), or glum (G). If he is cheerful today, then he will be C, S, or G tomorrow with respective probabilities 0.5, 0.4, 0.1. If he is feeling so-so today, then he will be C, S, or G tomorrow with probabilities 0.3, 0.4, 0.3. If he is glum today, then he will be C, S, or G tomorrow with probabilities 0.2, 0.3, 0.5.

Letting X_n denote Gary's mood on the nth day, then $\{X_n, n \geqslant 0\}$ is a three-state Markov chain (state $0 = C$, state $1 = S$, state $2 = G$) with transition probability matrix

$$\mathbf{P} = \begin{Vmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{Vmatrix}$$

■

Example 4.4 (Transforming a Process into a Markov Chain) Suppose that whether or not it rains today depends on previous weather conditions through the last two days. Specifically, suppose that if it has rained for the past two days, then it will rain tomorrow with probability 0.7; if it rained today but not yesterday, then it will rain tomorrow with probability 0.5; if it rained yesterday but not today, then it will rain tomorrow with probability 0.4; if it has not rained in the past two days, then it will rain tomorrow with probability 0.2.

If we let the state at time n depend only on whether or not it is raining at time n, then the preceding model is not a Markov chain (why not?). However, we can transform this model into a Markov chain by saying that the state at any time is determined by the weather conditions during both that day and the previous day. In other words, we can say that the process is in

state 0	if it rained both today and yesterday,	0.7
state 1	if it rained today but not yesterday,	0.5
state 2	if it rained yesterday but not today,	0.4
state 3	if it did not rain either yesterday or today.	0

The preceding would then represent a four-state Markov chain having a transition probability matrix

$$\mathbf{P} = \begin{Vmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{Vmatrix}$$

You should carefully check the matrix $\mathbf{P}$, and make sure you understand how it was obtained. ■

Example 4.5 (A Random Walk Model) A Markov chain whose state space is given by the integers $i = 0, \pm 1, \pm 2, \ldots$ is said to be a *random walk* if, for some number $0 < p < 1$,

$$P_{i,i+1} = p = 1 - P_{i,i-1}, \qquad i = 0, \pm 1, \ldots$$

The preceding Markov chain is called a random walk for we may think of it as being a model for an individual walking on a straight line who at each point of time either takes one step to the right with probability p or one step to the left with probability $1 - p$. ■

Example 4.6 (A Gambling Model) Consider a gambler who, at each play of the game, either wins \$1 with probability p or loses \$1 with probability $1 - p$. If we

suppose that our gambler quits playing either when he goes broke or he attains a fortune of \$$N$, then the gambler's fortune is a Markov chain having transition probabilities

$$P_{i,i+1} = p = 1 - P_{i,i-1}, \qquad i = 1, 2, \ldots, N-1,$$
$$P_{00} = P_{NN} = 1$$

States 0 and N are called *absorbing* states since once entered they are never left. Note that the preceding is a finite state random walk with absorbing barriers (states 0 and N). ■

Example 4.7 In most of Europe and Asia annual automobile insurance premiums are determined by use of a Bonus Malus (Latin for Good-Bad) system. Each policyholder is given a positive integer valued state and the annual premium is a function of this state (along, of course, with the type of car being insured and the level of insurance). A policyholder's state changes from year to year in response to the number of claims made by that policyholder. Because lower numbered states correspond to lower annual premiums, a policyholder's state will usually decrease if he or she had no claims in the preceding year, and will generally increase if he or she had at least one claim. (Thus, no claims is good and typically results in a decreased premium, while claims are bad and typically result in a higher premium.)

For a given Bonus Malus system, let $s_i(k)$ denote the next state of a policyholder who was in state i in the previous year and who made a total of k claims in that year. If we suppose that the number of yearly claims made by a particular policyholder is a Poisson random variable with parameter λ, then the successive states of this policyholder will constitute a Markov chain with transition probabilities

$$P_{i,j} = \sum_{k:s_i(k)=j} e^{-\lambda}\frac{\lambda^k}{k!}, \quad j \geqslant 0$$

Whereas there are usually many states (20 or so is not atypical), the following table specifies a hypothetical Bonus Malus system having four states.

		Next state if			
State	**Annual Premium**	**0 claims**	**1 claim**	**2 claims**	**$\geqslant$ 3 claims**
1	200	1	2	3	4
2	250	1	3	4	4
3	400	2	4	4	4
4	600	3	4	4	4

Thus, for instance, the table indicates that $s_2(0) = 1$; $s_2(1) = 3$; $s_2(k) = 4, k \geqslant 2$. Consider a policyholder whose annual number of claims is a Poisson random variable with parameter λ. If a_k is the probability that such a policyholder makes k claims in a year, then

$$a_k = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k \geqslant 0$$

For the Bonus Malus system specified in the preceding table, the transition probability matrix of the successive states of this policyholder is

$$\mathbf{P} = \begin{Vmatrix} a_0 & a_1 & a_2 & 1 - a_0 - a_1 - a_2 \\ a_0 & 0 & a_1 & 1 - a_0 - a_1 \\ 0 & a_0 & 0 & 1 - a_0 \\ 0 & 0 & a_0 & 1 - a_0 \end{Vmatrix} \qquad \blacksquare$$

4.2 Chapman–Kolmogorov Equations

We have already defined the one-step transition probabilities P_{ij}. We now define the n-step transition probabilities P_{ij}^n to be the probability that a process in state i will be in state j after n additional transitions. That is,

$$P_{ij}^n = P\{X_{n+k} = j | X_k = i\}, \quad n \geqslant 0, \ i, j \geqslant 0$$

Of course $P_{ij}^1 = P_{ij}$. The *Chapman–Kolmogorov equations* provide a method for computing these n-step transition probabilities. These equations are

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m \quad \text{for all } n, \ m \geqslant 0, \ \text{all } i, j \tag{4.2}$$

and are most easily understood by noting that $P_{ik}^n P_{kj}^m$ represents the probability that starting in i the process will go to state j in $n + m$ transitions through a path which takes it into state k at the nth transition. Hence, summing over all intermediate states k yields the probability that the process will be in state j after $n+m$ transitions. Formally, we have

$$\begin{aligned} P_{ij}^{n+m} &= P\{X_{n+m} = j | X_0 = i\} \\ &= \sum_{k=0}^{\infty} P\{X_{n+m} = j, X_n = k | X_0 = i\} \\ &= \sum_{k=0}^{\infty} P\{X_{n+m} = j | X_n = k, X_0 = i\} P\{X_n = k | X_0 = i\} \\ &= \sum_{k=0}^{\infty} P_{kj}^m P_{ik}^n \end{aligned}$$

If we let $\mathbf{P}^{(n)}$ denote the matrix of n-step transition probabilities P_{ij}^n, then Equation (4.2) asserts that

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

where the dot represents matrix multiplication.* Hence, in particular,

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

and by induction

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

That is, the n-step transition matrix may be obtained by multiplying the matrix $\mathbf{P}$ by itself n times.

Example 4.8 Consider Example 4.1 in which the weather is considered as a two-state Markov chain. If $\alpha = 0.7$ and $\beta = 0.4$, then calculate the probability that it will rain four days from today given that it is raining today.

Solution: The one-step transition probability matrix is given by

$$\mathbf{P} = \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix}$$

Hence,

$$\begin{aligned} \mathbf{P}^{(2)} = \mathbf{P}^2 &= \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix} \cdot \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix} \\ &= \begin{Vmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{Vmatrix}, \\ \mathbf{P}^{(4)} = (\mathbf{P}^2)^2 &= \begin{Vmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{Vmatrix} \cdot \begin{Vmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{Vmatrix} \\ &= \begin{Vmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{Vmatrix} \end{aligned}$$

and the desired probability P_{00}^4 equals 0.5749. ■

Example 4.9 Consider Example 4.4. Given that it rained on Monday and Tuesday, what is the probability that it will rain on Thursday?

Solution: The two-step transition matrix is given by

$$\begin{aligned} \mathbf{P}^{(2)} = \mathbf{P}^2 &= \begin{Vmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{Vmatrix} \cdot \begin{Vmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{Vmatrix} \\ &= \begin{Vmatrix} 0.49 & 0.12 & 0.21 & 0.18 \\ 0.35 & 0.20 & 0.15 & 0.30 \\ 0.20 & 0.12 & 0.20 & 0.48 \\ 0.10 & 0.16 & 0.10 & 0.64 \end{Vmatrix} \end{aligned}$$

* If $\mathbf{A}$ is an $N \times M$ matrix whose element in the ith row and jth column is a_{ij} and $\mathbf{B}$ is an $M \times K$ matrix whose element in the ith row and jth column is b_{ij}, then $\mathbf{A} \cdot \mathbf{B}$ is defined to be the $N \times K$ matrix whose element in the ith row and jth column is $\sum_{k=1}^{M} a_{ik} b_{kj}$.

Since rain on Thursday is equivalent to the process being in either state 0 or state 1 on Thursday, the desired probability is given by $P_{00}^2 + P_{01}^2 = 0.49 + 0.12 = 0.61$. ■

Example 4.10 An urn always contains 2 balls. Ball colors are red and blue. At each stage a ball is randomly chosen and then replaced by a new ball, which with probability 0.8 is the same color, and with probability 0.2 is the opposite color, as the ball it replaces. If initially both balls are red, find the probability that the fifth ball selected is red.

Solution: To find the desired probability we first define an appropriate Markov chain. This can be accomplished by noting that the probability that a selection is red is determined by the composition of the urn at the time of the selection. So, let us define X_n to be the number of red balls in the urn after the nth selection and subsequent replacement. Then $X_n, n \geqslant 0$, is a Markov chain with states 0, 1, 2 and with transition probability matrix $\mathbf{P}$ given by

$$\begin{pmatrix} 0.8 & 0.2 & 0 \\ 0.1 & 0.8 & 0.1 \\ 0 & 0.2 & 0.8 \end{pmatrix}$$

To understand the preceding, consider for instance $P_{1,0}$. Now, to go from 1 red ball in the urn to 0 red balls, the ball chosen must be red (which occurs with probability 0.5) and it must then be replaced by a ball of opposite color (which occurs with probability 0.2), showing that

$$P_{1,0} = (0.5)(0.2) = 0.1$$

To determine the probability that the fifth selection is red, condition on the number of red balls in the urn after the fourth selection. This yields

$$\begin{aligned} &P(\text{fifth selection is red}) \\ &\quad = \sum_{i=0}^{2} P(\text{fifth selection is red} | X_4 = i) P(X_4 = i | X_0 = 2) \\ &\quad = (0) P_{2,0}^4 + (0.5) P_{2,1}^4 + (1) P_{2,2}^4 \\ &\quad = 0.5 P_{2,1}^4 + P_{2,2}^4 \end{aligned}$$

To calculate the preceding we compute $\mathbf{P}^4$. Doing so yields

$$P_{2,1}^4 = 0.4352, \quad P_{2,2}^4 = 0.4872$$

giving the answer $P(\text{fifth selection is red}) = 0.7048$. ■

Example 4.11 Suppose that balls are successively distributed among 8 urns, with each ball being equally likely to be put in any of these urns. What is the probability that there will be exactly 3 nonempty urns after 9 balls have been distributed?

Solution: If we let X_n be the number of nonempty urns after n balls have been distributed, then $X_n, n \geqslant 0$ is a Markov chain with states $0, 1, \ldots, 8$ and transition probabilities

$$P_{i,i} = i/8 = 1 - P_{i,i+1}, \quad i = 0, 1, \ldots, 8$$

The desired probability is $P_{0,3}^9 = P_{1,3}^8$, where the equality follows because $P_{0,1} = 1$. Now, starting with 1 occupied urn, if we had wanted to determine the entire probability distribution of the number of occupied urns after 8 additional balls had been distributed we would need to consider the transition probability matrix with states $1, 2, \ldots, 8$. However, because we only require the probability, starting with a single occupied urn, that there are 3 occupied urns after an additional 8 balls have been distributed we can make use of the fact that the state of the Markov chain cannot decrease to collapse all states $4, 5, \ldots, 8$ into a single state 4 with the interpretation that the state is 4 whenever four or more of the urns are occupied. Consequently, we need only determine the eight-step transition probability $P_{1,3}^8$ of the Markov chain with states 1, 2, 3, 4 having transition probability matrix $\mathbf{P}$ given by

$$\begin{pmatrix} 1/8 & 7/8 & 0 & 0 \\ 0 & 2/8 & 6/8 & 0 \\ 0 & 0 & 3/8 & 5/8 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Raising the preceding matrix to the power 4 yields the matrix $\mathbf{P}^4$ given by

$$\begin{pmatrix} 0.0002 & 0.0256 & 0.2563 & 0.7178 \\ 0 & 0.0039 & 0.0952 & 0.9009 \\ 0 & 0 & 0.0198 & 0.9802 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Hence,

$$\begin{aligned} P_{1,3}^8 &= 0.0002 \times 0.2563 + 0.0256 \times 0.0952 + 0.2563 \times 0.0198 \\ &\quad + 0.7178 \times 0 = 0.00756 \end{aligned}$$

■

Consider a Markov chain with transition probabilities P_{ij}. Let $\mathscr{A}$ be a set of states, and suppose we are interested in the probability that the Markov chain ever enters any of the states in $\mathscr{A}$ by time m. That is, for a given state $i \notin \mathscr{A}$, we are interested in determining

$$\beta = P(X_k \in \mathscr{A} \text{ for some } k = 1, \ldots, m | X_0 = i)$$

To determine the preceding probability we will define a Markov chain $\{W_n, n \geqslant 0\}$ whose states are the states that are not in $\mathscr{A}$ plus an additional state, which we will call A in our general discussion (though in specific examples we will usually give it a different name). Once the $\{W_n\}$ Markov chain enters state A it remains there forever.

The new Markov chain is defined as follows. Letting X_n denote the state at time n of the Markov chain with transition probabilities $P_{i,j}$, define

$$N = \min\{n : X_n \in \mathscr{A}\}$$

and let $N = \infty$ if $X_n \notin \mathscr{A}$ for all n. In words, N is the first time the Markov chain enters the set of states $\mathscr{A}$. Now, define

$$W_n = \begin{cases} X_n, & \text{if } n < N \\ A, & \text{if } n \geqslant N \end{cases}$$

So the state of the $\{W_n\}$ process is equal to the state of the original Markov chain up to the point when the original Markov chain enters a state in $\mathscr{A}$. At that time the new process goes to state A and remains there forever. From this description it follows that $W_n, n \geqslant 0$ is a Markov chain with states $i, i \notin \mathscr{A}, A$ and with transition probabilities $Q_{i,j}$, given by

$$\begin{aligned} Q_{i,j} &= P_{i,j}, \quad \text{if} \quad i \notin \mathscr{A}, j \notin \mathscr{A} \\ Q_{i,A} &= \sum_{j \in \mathscr{A}} P_{i,j}, \quad \text{if} \quad i \notin \mathscr{A} \\ Q_{A,A} &= 1 \end{aligned}$$

Because the original Markov chain will have entered a state in $\mathscr{A}$ by time m if and only if the state at time m of the new Markov chain is A, we see that

$$\begin{aligned} &P(X_k \in \mathscr{A} \text{ for some } k = 1, \ldots, m | X_0 = i) \\ &\quad = P(W_m = A | X_0 = i) = P(W_m = A | W_0 = i) = Q_{i,A}^m \end{aligned}$$

That is, the desired probability is equal to an m-step transition probability of the new chain.

Example 4.12 In a sequence of independent flips of a fair coin, let N denote the number of flips until there is a run of three consecutive heads. Find

(a) $P(N \leqslant 8)$ and
(b) $P(N = 8)$.

Solution: To determine $P(N \leqslant 8)$, define a Markov chain with states $0, 1, 2, 3$ where for $i < 3$ state i means that we currently are on a run of i consecutive heads, and where state 3 means that a run of three consecutive heads has already occurred. Thus, the transition probability matrix is

$$\mathbf{P} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where, for instance, the values for row 2 are obtained by noting that if we currently are on a run of size 1 then the next state will be 0 if the next flip is a tail, or 2 if it is a

head. Hence, $P_{10} = P_{1,2} = 1/2$. Because there would be a run of three consecutive heads within the first eight flips if and only if $X_8 = 3$, the desired probability is $P^8_{0,3}$. Squaring $\mathbf{P}$ to obtain $\mathbf{P}^2$, then squaring the result to obtain $\mathbf{P}^4$, and then squaring that matrix gives the result

$$\mathbf{P}^8 = \begin{pmatrix} 81/256 & 44/256 & 24/256 & 107/256 \\ 68/256 & 37/256 & 20/256 & 131/256 \\ 44/256 & 24/256 & 13/256 & 175/256 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Hence, the probability that there will be a run of three consecutive heads within the first eight flips is $107/256 \approx .4180$.
(b) One way to obtain $P(N = 8)$, the probability that it takes eight flips to obtain the first run of three consecutive heads, is to use that

$$P(N = 8) = P(N \leqslant 8) - P(N \leqslant 7) = P^8_{0.3} - P^7_{0.3}$$

Another way to determine $P(N = 8)$ is to consider a Markov chain with states 0, 1, 2, 3, 4 where, as before, for $i < 3$ state i means that we currently are on a run of i consecutive heads, state 3 means that the first run of size 3 has just occurred, and state 4 that a run of size 3 occurred in the past. That is, this Markov chain has transition probability matrix

$$\mathbf{Q} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

N will equal 8 if starting in state 0 the preceding Markov chain is in state 3 after eight transitions. That is, $P(N = 8) = Q^8_{0.3}$. ■

Suppose now that we want to compute the probability that the $\{X_n, n \geqslant 0\}$ chain, starting in state i, enters state j at time m without ever entering any of the states in $\mathscr{A}$, where neither i nor j is in $\mathscr{A}$. That is, for $i, j \notin \mathscr{A}$, we are interested in

$$\alpha = P(X_m = j, X_k \notin \mathscr{A}, k = 1, \ldots, m-1 | X_0 = i)$$

Noting that the event that $X_m = j, X_k \notin \mathscr{A}, k = 1, \ldots, m-1$ is equivalent to the event that $W_m = j$, it follows that for $i, j \notin \mathscr{A}$,

$$\begin{aligned} P(X_m = j, X_k \notin \mathscr{A}, k = 1, \ldots, m-1 | X_0 = i) &= P(W_m = j | X_0 = i) \\ &= P(W_m = j | W_0 = i) = Q^m_{i,j}. \end{aligned}$$

Example 4.13 Consider a Markov chain with states 1, 2, 3, 4, 5, and suppose that we want to compute

$$P(X_4 = 2, X_3 \leqslant 2, X_2 \leqslant 2, X_1 \leqslant 2 | X_0 = 1)$$

That is, we want the probability that, starting in state 1, the chain is in state 2 at time 4 and has never entered any of the states in the set $\mathscr{A} = \{3, 4, 5\}$.

To compute this probability all we need to know are the transition probabilities $P_{11}, P_{12}, P_{21}, P_{22}$. So, suppose that

$$P_{11} = 0.3 \quad P_{12} = 0.3$$
$$P_{21} = 0.1 \quad P_{22} = 0.2$$

Then we consider the Markov chain having states 1, 2, 3 (we are giving state A the name 3), and having the transition probability matrix $\mathbf{Q}$ as follows:

$$\begin{pmatrix} 0.3 & 0.3 & 0.4 \\ 0.1 & 0.2 & 0.7 \\ 0 & 0 & 1 \end{pmatrix}$$

The desired probability is Q_{12}^4. Raising $\mathbf{Q}$ to the power 4 yields the matrix

$$\begin{pmatrix} 0.0219 & 0.0285 & 0.9496 \\ 0.0095 & 0.0124 & 0.9781 \\ 0 & 0 & 1 \end{pmatrix}$$

Hence, the desired probability is $\alpha = 0.0285$. ■

When $i \notin \mathscr{A}$ but $j \in \mathscr{A}$ we can determine the probability

$$\alpha = P(X_m = j, X_k \notin \mathscr{A}, k = 1, \ldots, m-1 | X_0 = i)$$

as follows.

$$\begin{aligned}
\alpha &= \sum_{r \notin \mathscr{A}} P(X_m = j, X_{m-1} = r, X_k \notin \mathscr{A}, k = 1, \ldots, m-2 | X_0 = i) \\
&= \sum_{r \notin \mathscr{A}} P(X_m = j | X_{m-1} = r, X_k \notin \mathscr{A}, k = 1, \ldots, m-2, X_0 = i) \\
&\quad \times P(X_{m-1} = r, X_k \notin \mathscr{A}, k = 1, \ldots, m-2 | X_0 = i) \\
&= \sum_{r \notin \mathscr{A}} P_{r,j}\, P(X_{m-1} = r, X_k \notin \mathscr{A}, k = 1, \ldots, m-2 | X_0 = i) \\
&= \sum_{r \notin \mathscr{A}} P_{r,j}\, Q_{i,r}^{m-1}
\end{aligned}$$

Also, when $i \in \mathscr{A}$ we could determine

$$\alpha = P(X_m = j, X_k \notin \mathscr{A}, k = 1, \ldots, m-1 | X_0 = i)$$

by conditioning on the first transition to obtain

$$\begin{aligned}
\alpha &= \sum_{r \notin \mathscr{A}} P(X_m = j, X_k \notin \mathscr{A}, \\
&\qquad\qquad k = 1, \ldots, m-1 | X_0 = i, X_1 = r) P(X_1 = r | X_0 = i) \\
&= \sum_{r \notin \mathscr{A}} P(X_{m-1} = j, X_k \notin \mathscr{A}, k = 1, \ldots, m-2 | X_0 = r) P_{i,r}
\end{aligned}$$

For instance, if $i \in \mathscr{A}$, $j \notin \mathscr{A}$ then the preceding equation yields

$$P(X_m = j, X_k \notin \mathscr{A}, k = 1, \ldots, m-1 | X_0 = i) = \sum_{r \notin \mathscr{A}} Q_{r,j}^{m-1} P_{i,r}$$

We can also compute the conditional probability of X_n given that the chain starts in state i and has not entered any state in $\mathscr{A}$ by time n, as follows. For $i, j \notin \mathscr{A}$,

$$\begin{aligned} &P\{X_n = j | X_0 = i, X_k \notin \mathscr{A}, k = 1, \ldots, n\} \\ &= \frac{P\{X_n = j, X_k \notin \mathscr{A}, k = 1, \ldots, n | X_0 = i\}}{P\{X_k \notin \mathscr{A}, k = 1, \ldots, n | X_0 = i\}} = \frac{Q_{i,j}^n}{\sum_{r \notin \mathscr{A}} Q_{i,r}^n} \end{aligned}$$

Remark So far, all of the probabilities we have considered are conditional probabilities. For instance, P_{ij}^n is the probability that the state at time n is j given that the initial state at time 0 is i. If the unconditional distribution of the state at time n is desired, it is necessary to specify the probability distribution of the initial state. Let us denote this by

$$\alpha_i \equiv P\{X_0 = i\}, \quad i \geqslant 0 \quad \left(\sum_{i=0}^{\infty} \alpha_i = 1\right)$$

All unconditional probabilities may be computed by conditioning on the initial state. That is,

$$\begin{aligned} P\{X_n = j\} &= \sum_{i=0}^{\infty} P\{X_n = j | X_0 = i\} P\{X_0 = i\} \\ &= \sum_{i=0}^{\infty} P_{ij}^n \alpha_i \end{aligned}$$

For instance, if $\alpha_0 = 0.4$, $\alpha_1 = 0.6$, in Example 4.8, then the (unconditional) probability that it will rain four days after we begin keeping weather records is

$$\begin{aligned} P\{X_4 = 0\} &= 0.4 P_{00}^4 + 0.6 P_{10}^4 \\ &= (0.4)(0.5749) + (0.6)(0.5668) \\ &= 0.5700 \end{aligned}$$

4.3 Classification of States

State j is said to be *accessible* from state i if $P_{ij}^n > 0$ for some $n \geqslant 0$. Note that this implies that state j is accessible from state i if and only if, starting in i, it is possible that the process will ever enter state j. This is true since if j is not accessible

from i, then

$$\begin{aligned} P\{\text{ever be in } j|\text{start in } i\} &= P\left\{\bigcup_{n=0}^{\infty}\{X_n = j\}\,\middle|\, X_0 = i\right\} \\ &\leqslant \sum_{n=0}^{\infty} P\{X_n = j|X_0 = i\} \\ &= \sum_{n=0}^{\infty} P_{ij}^n \\ &= 0 \end{aligned}$$

Two states i and j that are accessible to each other are said to *communicate*, and we write $i \leftrightarrow j$.

Note that any state communicates with itself since, by definition,

$$P_{ii}^0 = P\{X_0 = i|X_0 = i\} = 1$$

The relation of communication satisfies the following three properties:

(i) State i communicates with state i, all $i \geqslant 0$.
(ii) If state i communicates with state j, then state j communicates with state i.
(iii) If state i communicates with state j, and state j communicates with state k, then state i communicates with state k.

Properties (i) and (ii) follow immediately from the definition of communication. To prove (iii) suppose that i communicates with j, and j communicates with k. Thus, there exist integers n and m such that $P_{ij}^n > 0$, $P_{jk}^m > 0$. Now by the Chapman–Kolmogorov equations, we have

$$P_{ik}^{n+m} = \sum_{r=0}^{\infty} P_{ir}^n P_{rk}^m \geqslant P_{ij}^n P_{jk}^m > 0$$

Hence, state k is accessible from state i. Similarly, we can show that state i is accessible from state k. Hence, states i and k communicate.

Two states that communicate are said to be in the same *class*. It is an easy consequence of (i), (ii), and (iii) that any two classes of states are either identical or disjoint. In other words, the concept of communication divides the state space up into a number of separate classes. The Markov chain is said to be *irreducible* if there is only one class, that is, if all states communicate with each other.

Example 4.14 Consider the Markov chain consisting of the three states 0, 1, 2 and having transition probability matrix

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

It is easy to verify that this Markov chain is irreducible. For example, it is possible to go from state 0 to state 2 since

$$0 \to 1 \to 2$$

That is, one way of getting from state 0 to state 2 is to go from state 0 to state 1 (with probability $\frac{1}{2}$) and then go from state 1 to state 2 (with probability $\frac{1}{4}$). ■

Example 4.15 Consider a Markov chain consisting of the four states 0, 1, 2, 3 and having transition probability matrix

$$\mathbf{P} = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 \end{Vmatrix}$$

No

The classes of this Markov chain are $\{0, 1\}$, $\{2\}$, and $\{3\}$. Note that while state 0 (or 1) is accessible from state 2, the reverse is not true. Since state 3 is an absorbing state, that is, $P_{33} = 1$, no other state is accessible from it. ■

For any state i we let f_i denote the probability that, starting in state i, the process will ever reenter state i. State i is said to be *recurrent* if $f_i = 1$ and *transient* if $f_i < 1$.

Suppose that the process starts in state i and i is recurrent. Hence, with probability 1, the process will eventually reenter state i. However, by the definition of a Markov chain, it follows that the process will be starting over again when it reenters state i and, therefore, state i will eventually be visited again. Continual repetition of this argument leads to the conclusion that *if state i is recurrent then, starting in state i, the process will reenter state i again and again and again—in fact, infinitely often.*

On the other hand, suppose that state i is transient. Hence, each time the process enters state i there will be a positive probability, namely, $1 - f_i$, that it will never again enter that state. Therefore, starting in state i, the probability that the process will be in state i for exactly n time periods equals $f_i^{n-1}(1 - f_i)$, $n \geqslant 1$. In other words, *if state i is transient then, starting in state i, the number of time periods that the process will be in state i has a geometric distribution with finite mean* $1/(1 - f_i)$. 1/p

From the preceding two paragraphs, it follows that *state i is recurrent if and only if, starting in state i, the expected number of time periods that the process is in state i is infinite.* But, letting

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

we have that $\sum_{n=0}^{\infty} I_n$ represents the number of periods that the process is in state i. Also,

$$E\left[\sum_{n=0}^{\infty} I_n | X_0 = i\right] = \sum_{n=0}^{\infty} E[I_n | X_0 = i]$$

$$= \sum_{n=0}^{\infty} P\{X_n = i | X_0 = i\}$$

$$= \sum_{n=0}^{\infty} P_{ii}^n$$

We have thus proven the following.

Proposition 4.1 State i is

recurrent if $\sum_{n=1}^{\infty} P_{ii}^n = \infty$, no. of times

transient if $\sum_{n=1}^{\infty} P_{ii}^n < \infty$ finite

The argument leading to the preceding proposition is doubly important because it also shows that a transient state will only be visited a finite number of times (hence the name transient). This leads to the conclusion that in a finite-state Markov chain not all states can be transient. To see this, suppose the states are $0, 1, \ldots, M$ and suppose that they are all transient. Then after a finite amount of time (say, after time T_0) state 0 will never be visited, and after a time (say, T_1) state 1 will never be visited, and after a time (say, T_2) state 2 will never be visited, and so on. Thus, after a finite time $T = \max\{T_0, T_1, \ldots, T_M\}$ no states will be visited. But as the process must be in some state after time T we arrive at a contradiction, which shows that at least one of the states must be recurrent.

Another use of Proposition 4.1 is that it enables us to show that recurrence is a class property.

Corollary 4.2 If state i is recurrent, and state i communicates with state j, then state j is recurrent.

Proof. To prove this we first note that, since state i communicates with state j, there exist integers k and m such that $P_{ij}^k > 0$, $P_{ji}^m > 0$. Now, for any integer n

$$P_{jj}^{m+n+k} \geqslant P_{ji}^m P_{ii}^n P_{ij}^k$$

This follows since the left side of the preceding is the probability of going from j to j in $m + n + k$ steps, while the right side is the probability of going from j to j in $m + n + k$ steps via a path that goes from j to i in m steps, then from i to i in an additional n steps, then from i to j in an additional k steps.

From the preceding we obtain, by summing over n, that

$$\sum_{n=1}^{\infty} P_{jj}^{m+n+k} \geqslant P_{ji}^m P_{ij}^k \sum_{n=1}^{\infty} P_{ii}^n = \infty$$

since $P_{ji}^m P_{ij}^k > 0$ and $\sum_{n=1}^{\infty} P_{ii}^n$ is infinite since state i is recurrent. Thus, by Proposition 4.1 it follows that state j is also recurrent. ■

Remarks

(i) Corollary 4.2 also implies that transience is a class property. For if state i is transient and communicates with state j, then state j must also be transient. For if j were recurrent then, by Corollary 4.2, i would also be recurrent and hence could not be transient.

(ii) Corollary 4.2 along with our previous result that not all states in a finite Markov chain can be transient leads to the conclusion that all states of a finite irreducible Markov chain are recurrent.

Example 4.16 Let the Markov chain consisting of the states 0, 1, 2, 3 have the transition probability matrix

$$\mathbf{P} = \begin{Vmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{Vmatrix}$$

Determine which states are transient and which are recurrent.

irreducible

Solution: It is a simple matter to check that all states communicate and, hence, since this is a finite chain, all states must be recurrent. ■

Example 4.17 Consider the Markov chain having states 0, 1, 2, 3, 4 and

$$\mathbf{P} = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{2} \end{Vmatrix}$$

Till example 4.17

Determine the recurrent state.

Solution: This chain consists of the three classes {0, 1}, {2, 3}, and {4}. The first two classes are recurrent and the third transient. ■

Example 4.18 (A Random Walk) Consider a Markov chain whose state space consists of the integers $i = 0, \pm 1, \pm 2, \ldots$, and has transition probabilities given by

$$P_{i,i+1} = p = 1 - P_{i,i-1}, \quad i = 0, \pm 1, \pm 2, \ldots$$

where $0 < p < 1$. In other words, on each transition the process either moves one step to the right (with probability p) or one step to the left (with probability $1 - p$). One colorful interpretation of this process is that it represents the wanderings of a drunken man as he walks along a straight line. Another is that it represents the winnings of a gambler who on each play of the game either wins or loses one dollar.

4.4 Long-Run Proportions and Limiting Probabilities

For pairs of states $i \neq j$, let $f_{i,j}$ denote the probability that the Markov chain, starting in state i, will ever make a transition into state j. That is,

$$f_{i,j} = P(X_n = j \text{ for some } n > 0 | X_0 = i)$$

We then have the following result.

Proposition 4.3 If i is recurrent and i communicates with j, then $f_{i,j} = 1$.

Proof. Because i and j communicate there is a value n such that $P^n_{i,j} > 0$. Let $X_0 = i$ and say that the first opportunity is a success if $X_n = j$, and note that the first opportunity is a success with probability $P^n_{i,j} > 0$. If the first opportunity is not a success then consider the next time (after time n) that the chain enters state i. (Because state i is recurrent we can be certain that it will eventually reenter state i after time n.) Say that the second opportunity is a success if n time periods later the Markov chain is in state j. If the second opportunity is not a success then wait until the next time the chain enters state i and say that the third opportunity is a success if n time periods later the Markov chain is in state j. Continuing in this manner, we can define an unlimited number of opportunities, each of which is a success with the same positive probability $P^n_{i,j}$. Because the number of opportunities until the first success occurs is geometric with parameter $P^n_{i,j}$, it follows that with probability 1 a success will eventually occur and so, with probability 1, state j will eventually be entered. ■

If state j is recurrent, let m_j denote the expected number of transitions that it takes the Markov chain when starting in state j to return to that state. That is, with

$$N_j = \min\{n > 0 : X_n = j\}$$

equal to the number of transitions until the Markov chain makes a transition into state j,

$$m_j = E[N_j | X_0 = j]$$

finite

Definition: Say that the recurrent state j is *positive recurrent* if $m_j < \infty$ and say that it is *null recurrent* if $m_j = \infty$. never come back

Now suppose that the Markov chain is irreducible and recurrent. In this case we now show that the long-run proportion of time that the chain spends in state j is equal to $1/m_j$. That is, letting π_j denote the long-run proportion of time that the Markov chain is in state j, we have the following proposition.

Proposition 4.4 If the Markov chain is irreducible and recurrent, then for any initial state

$$\pi_j = 1/m_j$$

Proof. Suppose that the Markov chain starts in state i, and let T_1 denote the number of transitions until the chain enters state j; then let T_2 denote the additional number of

transitions from time T_1 until the Markov chain next enters state j; then let T_3 denote the additional number of transitions from time $T_1 + T_2$ until the Markov chain next enters state j, and so on. Note that T_1 is finite because Proposition 4.3 tells us that with probability 1 a transition into j will eventually occur. Also, for $n \geqslant 2$, because T_n is the number of transitions between the $(n-1)$th and the nth transition into state j, it follows from the Markovian property that $T_2, T_3, \ldots$ are independent and identically distributed with mean m_j. Because the nth transition into state j occurs at time $T_1 + \ldots + T_n$ we obtain that π_j, the long-run proportion of time that the chain is in state j, is

$$\begin{aligned}
\pi_j &= \lim_{n\to\infty} \frac{n}{\sum_{i=1}^n T_i} \\
&= \lim_{n\to\infty} \frac{1}{\frac{1}{n}\sum_{i=1}^n T_i} \\
&= \lim_{n\to\infty} \frac{1}{\frac{T_1}{n} + \frac{T_2+\ldots+T_n}{n}} \\
&= \frac{1}{m_j}
\end{aligned}$$

where the last equality follows because $\lim_{n\to\infty} T_1/n = 0$ and, from the strong law of large numbers, $\lim_{n\to\infty} \frac{T_2+\ldots+T_n}{n} = \lim_{n\to\infty} \frac{T_2+\ldots+T_n}{n-1}\,\frac{n-1}{n} = m_j$. ■

Because $m_j < \infty$ is equivalent to $1/m_j > 0$, it follows from the preceding that state j is positive recurrent if and only if $\pi_j > 0$. We now exploit this to show that positive recurrence is a class property.

Proposition 4.5 If i is positive recurrent and $i \leftrightarrow j$ then j is positive recurrent.

Proof. Suppose that i is positive recurrent and that $i \leftrightarrow j$. Now, let n be such that $P_{i,j}^n > 0$. Because π_i is the long-run proportion of time that the chain is in state i, and $P_{i,j}^n$ is the long-run proportion of time when the Markov is in state i that it will be in state j after n transitions

$$\begin{aligned}
\pi_i P_{i,j}^n &= \text{long-run proportion of time the chain is in } i \\
&\quad\ \text{and will be in } j \text{ after } n \text{ transitions} \\
&= \text{long-run proportion of time the chain is in } j \\
&\quad\ \text{and was in } i \ n \text{ transitions ago} \\
&\leqslant \text{long-run proportion of time the chain is in } j
\end{aligned}$$

Hence, $\pi_j \geqslant \pi_i P_{i,j}^n > 0$, showing that j is positive recurrent. ■

Remarks

(i) It follows from the preceding result that null recurrence is also a class property. For suppose that i is null recurrent and $i \leftrightarrow j$. Because i is recurrent and $i \leftrightarrow j$ we can conclude that j is recurrent. But if j were positive recurrent then by the preceding proposition i would also be positive recurrent. Because i is not positive recurrent, neither is j.

(ii) An irreducible finite state Markov chain must be positive recurrent. For we know that such a chain must be recurrent; hence, all its states are either positive recurrent or null recurrent. If they were null recurrent then all the long run proportions would equal 0, which is impossible when there are only a finite number of states. Consequently, we can conclude that the chain is positive recurrent. ■

To determine the long-run proportions $\{\pi_j, j \geqslant 1\}$, note, because π_i is the long-run proportion of transitions that come from state i, that

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

Summing the preceding over all i now yields that

$$\pi_j = \sum_i \pi_i P_{i,j}$$

Indeed, the following important theorem can be proven.

Theorem 4.1 Consider an irreducible Markov chain. If the chain is positive recurrent then the long-run proportions are the unique solution of the equations

$$\pi_j = \sum_i \pi_i P_{i,j}, \quad j \geqslant 1$$

$$\sum_j \pi_j = 1$$

Moreover, if there is no solution of the preceding linear equations, then the Markov chain is either transient or null recurrent and all $\pi_j = 0$.

Example 4.20 Consider Example 4.1, in which we assume that if it rains today, then it will rain tomorrow with probability α; and if it does not rain today, then it will rain tomorrow with probability β. If we say that the state is 0 when it rains and 1 when it does not rain, then by Theorem 4.1 the long-run proportions π_0 and π_1 are given by

$$\pi_0 = \alpha\pi_0 + \beta\pi_1,$$
$$\pi_1 = (1-\alpha)\pi_0 + (1-\beta)\pi_1,$$
$$\pi_0 + \pi_1 = 1$$

which yields that

$$\pi_0 = \frac{\beta}{1+\beta-\alpha}, \quad \pi_1 = \frac{1-\alpha}{1+\beta-\alpha}$$

For example if $\alpha = 0.7$ and $\beta = 0.4$, then the long-run proportion of rain is $\pi_0 = \frac{4}{7} = 0.571$. ■

Example 4.21 Consider Example 4.3 in which the mood of an individual is considered as a three-state Markov chain having a transition probability matrix

$$\mathbf{P} = \begin{Vmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{Vmatrix}$$

In the long run, what proportion of time is the process in each of the three states?

Solution: The long run proportions π_i, $i = 0, 1, 2$, are obtained by solving the set of equations in Equation (4.1). In this case these equations are

$$\begin{aligned}
\pi_0 &= 0.5\pi_0 + 0.3\pi_1 + 0.2\pi_2, \\
\pi_1 &= 0.4\pi_0 + 0.4\pi_1 + 0.3\pi_2, \\
\pi_2 &= 0.1\pi_0 + 0.3\pi_1 + 0.5\pi_2, \\
\pi_0 + \pi_1 + \pi_2 &= 1
\end{aligned}$$

Solving yields

$$\pi_0 = \tfrac{21}{62}, \quad \pi_1 = \tfrac{23}{62}, \quad \pi_2 = \tfrac{18}{62}$$ ■

Example 4.22 (A Model of Class Mobility) A problem of interest to sociologists is to determine the proportion of society that has an upper- or lower-class occupation. One possible mathematical model would be to assume that transitions between social classes of the successive generations in a family can be regarded as transitions of a Markov chain. That is, we assume that the occupation of a child depends only on his or her parent's occupation. Let us suppose that such a model is appropriate and that the transition probability matrix is given by

$$\mathbf{P} = \begin{Vmatrix} 0.45 & 0.48 & 0.07 \\ 0.05 & 0.70 & 0.25 \\ 0.01 & 0.50 & 0.49 \end{Vmatrix} \tag{4.8}$$

Till Example 4.22

That is, for instance, we suppose that the child of a middle-class worker will attain an upper-, middle-, or lower-class occupation with respective probabilities 0.05, 0.70, 0.25.

The long-run proportions π_i thus satisfy

$$\begin{aligned}
\pi_0 &= 0.45\pi_0 + 0.05\pi_1 + 0.01\pi_2, \\
\pi_1 &= 0.48\pi_0 + 0.70\pi_1 + 0.50\pi_2, \\
\pi_2 &= 0.07\pi_0 + 0.25\pi_1 + 0.49\pi_2, \\
\pi_0 + \pi_1 + \pi_2 &= 1
\end{aligned}$$

Hence,

$$\pi_0 = 0.07, \quad \pi_1 = 0.62, \quad \pi_2 = 0.31$$

In other words, a society in which social mobility between classes can be described by a Markov chain with transition probability matrix given by Equation (4.8) has, in the long run, 7 percent of its people in upper-class jobs, 62 percent of its people in middle-class jobs, and 31 percent in lower-class jobs. ■