

Book

Chapter 13

# Quantitative Methods for Decision Making in Business

## Markov Analysis

R-E  
Topic

Date

Page

## Markov Analysis

$P$  is a square matrix with nonnegative elements and row elements that sum to unity. Such a matrix is called a stochastic matrix. Any stochastic matrix can serve as a matrix of transition probabilities; together with an initial probability distribution of states it completely defines a Markov chain.

If the probability of reaching a particular state depends on the two previous states, we have a second-order Markov chain. We will consider only first-order Markov chains, since the topic of higher-order Markov chains is much more advanced.

We will now illustrate some of the basic concepts with an example of a Markov chain for a brand-switching problem, as it is generally referred to, with just two states.

## EXAMPLE 13.1

The Trustworthy Car Company (hereafter TCC), a well-established and enterprising dealership, has always maintained files on its customers' car purchases. For every customer, a record has been kept of the available details of all cars purchased, naturally including whether the car was bought from TCC or another source. Over the years, TCC's customers have exhibited a high degree of loyalty, as measured by the fact that customers owning a TCC car have purchased their next car from TCC 80 percent of the time. Also, former customers purchasing a car from another source have switched back to TCC for their next car 60 percent of the time.

Let state 1 be the condition of owning a TCC car and state 2 that of owning a non-TCC car. The transition probabilities are then

$$p_{11} = 0.8, \quad p_{12} = 0.2, \quad p_{21} = 0.6, \quad p_{22} = 0.4.$$

The transition probability matrix can be structured as in Table 13.1.

Table 13.1 Brand-Switching Transition Probabilities

Current State $i$	Next State $j$	
	$p_{ij}$	$p_{ji}$
1 (TCC car)	0.8	0.2
2 (non-TCC car)	0.6	0.4

The transition probabilities can also be represented by what is called a transition diagram, which was presented earlier in the derivations of queueing models in the appendices to Chapter 12. For this example, the transition diagram is shown as Figure 13.1. The nodes indicate states, and the directed line segments, branches, or arcs represent the allowable transitions between states. The transition probability is specified for each possible transition.

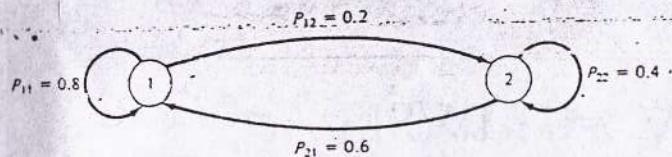


Figure 13.1 Brand-Switching Transition Diagram

We wish to investigate the probabilistic behavior of this Markov chain over a number of transitions.

Let  $s_{kj}(n)$  = probability the system occupies state  $j$  after  $n$  transitions, given that it started in state  $k$  at time 0.

$S(n)$  = the state row vector of probabilities after  $n$  transitions, with components  $s_{kj}(n)$ .

With  $N$  possible states,

$$\sum_{j=1}^N s_{kj}(n) = 1, \quad \text{for all } k. \quad (13.1)$$

At transition  $n + 1$ , the probability of being in state  $j$  is equal to the probability of being in state  $i$  on the previous transition weighted by the probability of a transition from state  $i$  to state  $j$ , for all possible values of  $i$ . Thus,

$$s_{kj}(n+1) = s_{k1}(n)p_{1j} + s_{k2}(n)p_{2j} + \cdots + s_{kN}(n)p_{Nj}.$$

Expressing this in a more convenient notational form,

$$s_{kj}(n+1) = \sum_{i=1}^N s_{ki}(n)p_{ij}, \quad n = 0, 1, 2, \dots \quad (13.2)$$

If we start with a customer's purchase of a TCC car,  $s_{11}(0) = 1$  and  $s_{12}(0) = 0$ , so

$$S(0) = [1 \quad 0].$$

After the first transition,

$$S(1) = S(0)P = [1 \quad 0] \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} = [0.5 \quad 0.2].$$

Continuing this procedure through three more transitions,

$$S(2) = S(1)P = [0.5 \quad 0.2] \begin{bmatrix} 0.5 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} = [0.76 \quad 0.24]$$

$$S(3) = S(2)P = [0.76 \quad 0.24] \begin{bmatrix} 0.5 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} = [0.752 \quad 0.248]$$

$$S(4) = S(3)P = [0.752 \quad 0.248] \begin{bmatrix} 0.5 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} = [0.7504 \quad 0.2496].$$

If we let  $s_{ij}(\infty)$  be the probability of being in state  $j$  after a large number of transitions starting in state  $i$  at time 0, it appears that the state probabilities are converging rapidly to limiting values  $s_{11}(\infty) = 0.75$  and  $s_{21}(\infty) = 0.25$ . A little later, we will show that this is indeed true.

Now, suppose that the initial state had been one in which a customer owned a non-TCC car, so that  $s_{21}(0) = 0$  and  $s_{22}(0) = 1$ , giving

$$S(0) = [0 \ 1].$$

The first four state row vectors are then

$$S(1) = [0 \ 1] \begin{bmatrix} 0.5 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} = [0.6 \ 0.4]$$

$$S(2) = [0.6 \ 0.4] \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} = [0.72 \ 0.28]$$

$$S(3) = [0.72 \ 0.28] \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} = [0.744 \ 0.256]$$

$$S(4) = [0.744 \ 0.256] \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} = [0.7488 \ 0.2512].$$

Again, we see that the limiting values of the state probabilities appear to be  $s_{11}(\infty) = 0.75$  and  $s_{21}(\infty) = 0.25$ .

As a third case, suppose we knew that a customer was about to purchase a new car, but we had no information about the present car. We could then assume a probability distribution for the initial state such that  $S(0) = [r \ 1-r]$ , denoting that the probability of being in state 1 is  $r$  and the probability of being in state 2 is the complementary probability  $1-r$ . Since  $s_{11}(\infty) = s_{21}(\infty) = 0.75$ , and  $s_{12}(\infty) = s_{22}(\infty) = 0.25$ , as we have seen previously, the limiting state probabilities are the same regardless of the initial state, so the value of the probability  $r$  has no effect on the limiting state probabilities. When the limiting state probabilities become independent of the initial state as  $n$  becomes relatively large, such a Markov process is called an ergodic process.

We can also calculate the state row vectors in a different way. Since  $S(2) = S(1)P$ , and  $S(1) = S(0)P$ ,

$$S(2) = S(1)P \doteq S(0)P \cdot P = S(0)P^2$$

and, in general,

$$S(n) = S(0)P^n. \quad (13.3)$$

$P^n$  is called the multistep transition probability matrix.

In this example,

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix}$$

$$P^2 = P \cdot P = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} = \begin{bmatrix} 0.76 & 0.24 \\ 0.72 & 0.28 \end{bmatrix}$$

$$P^3 = P \cdot P^2 = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} \begin{bmatrix} 0.76 & 0.24 \\ 0.72 & 0.28 \end{bmatrix} = \begin{bmatrix} 0.752 & 0.248 \\ 0.744 & 0.256 \end{bmatrix}$$

$$P^4 = P^2 \cdot P^2 = \begin{bmatrix} 0.76 & 0.24 \\ 0.72 & 0.28 \end{bmatrix} \begin{bmatrix} 0.76 & 0.24 \\ 0.72 & 0.28 \end{bmatrix} = \begin{bmatrix} 0.7504 & 0.2496 \\ 0.7488 & 0.2512 \end{bmatrix}.$$

It is apparent that in the limit, as  $n \rightarrow \infty$ ,

$$P^\infty = \begin{bmatrix} 0.75 & 0.25 \\ 0.75 & 0.25 \end{bmatrix}.$$

To find the state row vector  $S(3)$  for the first case, where  $S(0) = [1 \ 0]$ , we could have calculated

$$S(3) = S(0)P^3 = [1 \ 0] \begin{bmatrix} 0.752 & 0.248 \\ 0.744 & 0.256 \end{bmatrix} = [0.752 \ 0.248],$$

which is, of course, the same as the earlier result.

To summarize the results of the analysis, regardless of whether the customer starts with a TCC or a non-TCC car, you could give odds of approximately 3 to 1 that, after about three transitions (purchases of three cars), the next car would be a TCC car. Furthermore, over succeeding periods, 75 percent of the cars owned can be expected to be from TCC.

Before going on, we will now differentiate between the various types of states in a Markov chain.

Recurrent State State  $j$  is recurrent, or persistent, if the probability is 1 that the process will eventually return to state  $j$ , having started in state  $j$ .

In Example 13.1, assuming that a family (or their heirs) continue purchasing cars in the same fashion into the indefinite future, both states are recurrent. A purchase (at any point in time) of a TCC car will eventually be followed by another purchase of a TCC car, and, likewise, a purchase of a non-TCC car must at some time be followed by another non-TCC car purchase.

Transient State State  $j$  is transient, or nonrecurrent, if the probability is 1 that the process will eventually not return to state  $j$ , having started in state  $j$ .

Absorbing State State  $j$  is absorbing, or trapping, if  $p_{jj} = 1$ . Once the process reaches state  $j$ , it can never leave.

To illustrate these last two state types, consider a modification of Example 13.1. Suppose that a family starts by purchasing a TCC car, but when they purchase a non-TCC car for the first time, they never again go back to TCC. The transition diagram would then be as shown in Figure 13.2. Although the process may remain in state 1 for a number of transitions, state 1 now becomes a transient state, because once the process leaves state 1, it can never return. State 2 is an absorbing state. When the process reaches state 2, as it eventually must, it remains in state 2.

\* Irreducible Markov Chain: A Markov chain is irreducible if every state can be reached from every other state in a finite no. of steps.

We will have much more to say later about Markov processes with transient and absorbing states; but at this point let us return to the analysis of a process with only recurrent states.

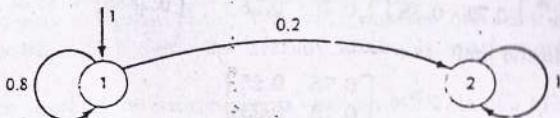


Figure 13.2 Brand-Switching Transition Diagram with Absorption

Example 13.1 concentrated on the transient behavior of a Markov process and indicated the limiting probability values as the number of transitions became large. Now, let us directly analyze the limiting, or steady-state, behavior of such a two-state process.

## 13.2 Steady-State Analysis

In many situations, we may be primarily interested in what occurs in the long run when the process achieves a steady-state condition. To show what occurs, suppose we work with a two-state Markov process which, if in state 1, has a probability  $p_{12} = a$  of moving to state 2. If in state 2, it has a probability  $p_{21} = b$  of moving to state 1. Therefore, the transition probability matrix  $P$  can be expressed as

$$P = \begin{bmatrix} 1 - a & a \\ b & 1 - b \end{bmatrix}$$

As we have shown in Example 13.1, as the number of transitions  $n$  becomes large, the state probability row vector  $S(n)$  reaches a limiting value  $S(\infty)$ , which we hereafter denote as just  $S$ . Then, the following relationship must apply:

$$S = SP \quad \text{for } n \text{ large.} \quad (13.4)$$

If we now denote  $s_1$  and  $s_2$  as the limiting state probabilities for a two-state process and use the preceding transition probability matrix, we can write

$$[s_1 \ s_2] = [s_1 \ s_2] \begin{bmatrix} 1 - a & a \\ b & 1 - b \end{bmatrix}$$

Separating this into individual equations,

$$\begin{aligned} s_1 &= (1 - a)s_1 + bs_2 \\ s_2 &= as_1 + (1 - b)s_2 \end{aligned}$$

Combining terms,

$$\begin{aligned} as_1 - bs_2 &= 0 \\ -as_1 + bs_2 &= 0. \end{aligned}$$

Because of the complementary probabilities in the transition matrix, these equations are dependent, so we have only one independent equation in two unknowns. If we solved that equation, we would obviously be able to determine only the ratio of  $s_1$  to  $s_2$ . To solve for  $s_1$  and  $s_2$  as probabilities, we drop one of the preceding equations and add the constraint that  $s_1 + s_2 = 1$  since the limiting state probabilities must sum to 1. The resultant equations are

$$\begin{aligned} as_1 - bs_2 &= 0 \\ s_1 + s_2 &= 1. \end{aligned}$$

Solving, we find that

$$s_1 = \frac{b}{a + b}, \quad s_2 = \frac{a}{a + b}$$

so that, for the general two-state Markov chain,

$$S = \left[ \frac{b}{a + b} \frac{a}{a + b} \right] = \left[ \frac{p_{21}}{p_{12} + p_{21}} \frac{p_{12}}{p_{12} + p_{21}} \right]. \quad (13.5)$$

We can also arrive at Equation (13.5) by working from the transition diagram of Figure 13.3 and applying the principle of what Howard (1971) terms the equilibrium of probabilistic flows. This procedure should lead to additional insights.

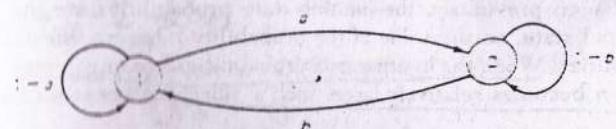


Figure 13.3 General Two-State Transition Diagram

In the steady state, the probability of a transition into state  $i$  must be the same as the probability of a transition out of state  $i$ . For example, the probability of leaving state 1 is equal to the probability  $s_1$  of being in state 1 multiplied by the probability  $a$  of leaving state 1. The probability of entering state 1 is the probability  $s_2$  of being in state 2 multiplied by the probability  $b$  of entering state 1. We can express this as

$$as_1 = bs_2.$$

## Markov Analysis

This equation, in combination with the requirement for  $s_1$  and  $s_2$  to sum to 1, give the same two equations used earlier to derive Equation (13.5).

The limiting multistep transition probability matrix will be of the following form for the general case.

$$P^\infty = \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a-b} & \frac{a}{a-b} \end{bmatrix} = \begin{bmatrix} \frac{p_{21}}{p_{12} + p_{21}} & \frac{p_{12}}{p_{12} + p_{21}} \\ \frac{p_{21}}{p_{12} - p_{21}} & \frac{p_{12}}{p_{12} - p_{21}} \end{bmatrix}. \quad (13.6)$$

This is so because, in the steady state, the following relationship must hold for any initial state probability vector  $S(0)$ .

$$S = S(0)P^\infty.$$

Let  $S(0)$  be of the form  $[r \ 1-r]$ , where  $0 \leq r \leq 1$ . Then,

$$\begin{aligned} S(0)P^\infty &= [r \ 1-r] \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a-b} \\ \frac{b}{a-b} & \frac{a}{a-b} \end{bmatrix} \\ &= \left[ \frac{b}{a+b} (r+1-r) \ \frac{a}{a-b} (r+1-r) \right] \\ &= \left[ \frac{b}{a+b} \ \frac{a}{a-b} \right] = S. \end{aligned}$$

independent of the value of  $r$ .

Going back to Example 13.1 to check the results, the one-step transition probability matrix was

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix}.$$

The limiting state probability vector would then be, from Equation 13.5:

$$S = \begin{bmatrix} 0.6 & 0.2 \\ 0.2 & 0.8 \end{bmatrix} = [0.75 \ 0.25].$$

The limiting multistep transition probability matrix, from Equation 13.6, would be

$$P^\infty = \begin{bmatrix} 0.75 & 0.25 \\ 0.75 & 0.25 \end{bmatrix}.$$

These answers, of course, are the same as the limiting values found in Example 13.1.

For a Markov chain with  $n$  recurrent states, the matrix relationship in Equation (13.4), with any one individual equation replaced by an equation where all the state probabilities are required to sum to 1, can be utilized to determine the steady-state probability vector. It is necessary only to solve a system of  $n$  simultaneous linear equations in  $n$  unknowns. Almost every computer software system has a canned program to do this with a minimum of effort on the user's part.

## EXAMPLE 13.2

In Los Angeles, there are three TV stations fighting it out for the top rating for the 6:00 P.M. one-hour newscast. At the end of each week, the leader is the station with the highest estimated average fraction of the viewing audience during that time slot. Over a period of time, data have been obtained on the relationship between ratings leadership in successive weekly periods. This information is presented in the form of a transition probability table, Table 13.2.

Table 13.2 Transition Probabilities for TV Station News Competition

This Week's Leader	Next Week's Leader		
	A	B	C
Station A	0.40	0.35	0.25
Station B	0.10	0.70	0.20
Station C	0.15	0.25	0.60

Over a long period of time, what proportion of the time will each station lead the weekly ratings?

A transition diagram will be helpful and is shown as Figure 13.4.

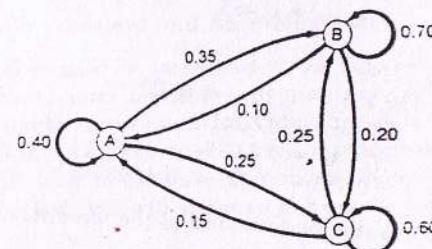


Figure 13.4 Transition Diagram, TV Ratings Problem

To determine the steady-state probabilities, we utilize the principle of the equilibrium of probabilistic flows, whereby the sum of the transition probabilities out of any state equals the sum of the transition probabilities into that state.

We then have:

$$\text{For state A: } (0.35 + 0.25)s_A = 0.10s_B + 0.15s_C.$$

$$\text{For state B: } (0.10 + 0.20)s_B = 0.35s_A - 0.25s_C.$$

$$\text{For state C: } (0.15 + 0.25)s_C = 0.25s_A + 0.20s_B.$$

Grouping terms, placing all variables on the left, and adding the constraint that the steady-state probabilities must sum to 1:

$$\begin{aligned} 0.60s_A - 0.10s_B - 0.15s_C &= 0 \\ -0.35s_A + 0.30s_B - 0.25s_C &= 0 \\ -0.25s_A - 0.20s_B + 0.40s_C &= 0 \\ s_A + s_B + s_C &= 1 \end{aligned}$$

One of the first three equations is redundant, so any one of those equations can be dropped. As a check on those three equations, observe that the sum of the coefficients of each  $s_i$  ( $i = A, B, C$  in this example) must be zero. Now, suppose we arbitrarily eliminate the third equation and then substitute (based on the fourth equation)  $s_A = 1 - s_B - s_C$  in the first two equations.

$$\begin{aligned} 0.60(1 - s_B - s_C) - 0.10s_B - 0.15s_C &= 0 \\ -0.35(1 - s_B - s_C) + 0.30s_B - 0.25s_C &= 0 \end{aligned}$$

Combining terms,

$$\begin{aligned} -0.70s_B - 0.75s_C &= -0.60 \\ 0.65s_B + 0.10s_C &= 0.35 \end{aligned}$$

Multiplying the second equation by 7.5 and adding the two equations,

$$\begin{aligned} 4.175s_B &= 2.025 \\ s_B &= 0.485 \end{aligned}$$

Solving for  $s_C$  and then  $s_A$ :

$$\begin{aligned} 0.65(0.485) + 0.10s_C &= 0.35 \\ 0.10s_C &= 0.0347 \\ s_C &= 0.347 \\ s_A &= 1 - 0.485 - 0.347 = 0.168 \end{aligned}$$

On the basis of these calculations, the performance of the three stations can be expected to be as follows over the long run:

Weekly Ratings Leader	Percentage of the Time
Station A	16.8
Station B	48.5
Station C	34.7

### 13.3 First-Passage Times

The number of transitions required to leave state  $i$  and to reach state  $j$  for the first time has come to be known as the *first-passage time*, a random variable. We will show how expected first-passage times can be determined and then will demonstrate

their applicability to problems involving recurrent states and also problems where there are absorbing states. Note carefully that, in the present context, time means "number of transitions."

#### Systems with Recurrent States

In the brand-switching problem of Example 13.1, the expected time between the purchase of a TCC car and the next purchase of a non-TCC car is just the expected first-passage time from state 1 to state 2. We will now show how such expected first-passage times can be calculated.

(Let  $\mu_{ij}$  = expected first-passage time (number of transitions) from state  $i$  to state  $j$ .

For systems containing only recurrent states, the value of  $\mu_{ii}$  can be expressed in the form of the following relationship.

$$\mu_{ii} = p_{ii} + \sum_{k \neq i} p_{ik}(1 - \mu_{kj}). \quad (13.7)$$

The first term, which is really  $p_{ii}$  multiplied by 1, expresses the fact that the transition from state  $i$  to state  $j$  will be made in one step with probability  $p_{ij}$ . Each term in the summation may be explained as follows: It takes one step to go from state  $i$  to state  $k$  ( $k \neq j$ ) and then an additional expected value of  $\mu_{kj}$  steps to get to state  $j$  for the first time. For each possible value of  $k$  (excluding  $k = j$ ), the expected number of steps of the path from state  $i$  to state  $k$  and then on to state  $j$  for the first time must then be weighted by the one-step transition probability  $p_{ik}$  of going from state  $i$  to state  $k$ . By breaking the summation in Equation (13.7) into two parts and then observing that all transition probabilities out of state  $i$  must sum to 1, a simpler result can be obtained.

$$\begin{aligned} \mu_{ii} &= p_{ii} + \sum_{k \neq j} p_{ik} + \sum_{k \neq i, k \neq j} p_{ik}\mu_{kj} \\ &= 1 + \sum_{k \neq j} p_{ik}\mu_{kj} \end{aligned} \quad (13.8)$$

To solve for the first-passage times to a given state  $j$  from all other states, this equation enables us to set up one equation for each of the other states. For a system with  $n$  states, the result is a set of  $n - 1$  linear equations in  $n - 1$  unknowns.

In Example 13.1, the expected first-passage time from state 1 to state 2 is the expected number of purchases between the purchase of a TCC car and a non-TCC car. From Equation (13.8),

$$\mu_{12} = 1 + p_{11}\mu_{12}, \quad \mu_{12} = \frac{1}{1 - p_{11}}$$

Since  $p_{11} = 0.8$ ,

$$\mu_{12} = \frac{1}{1 - 0.8} = 5$$

The expected time between the purchase of a TCC car and the next purchase of a non-TCC car is 5 purchases. Similarly,

$$\mu_{21} = 1 + p_{22}\mu_{21}, \quad \mu_{21} = \frac{1}{1 - p_{22}} = \frac{1}{1 - 0.4} = 1.67.$$

The expected number of purchases between purchase of a non-TCC car and the next purchase of a TCC car is 1.67.

In Equation (13.7), when  $j = i$ , we call the quantity  $\mu_{ii}$  the expected recurrence time, the expected number of steps between a start in state  $i$  and the next return to that same state. It is not difficult to show that the expected recurrence time  $\mu_{ii}$  is inversely proportional to the limiting state occupancy probability  $s_i$ , so that

$$\mu_{ii} = \frac{1}{s_i}. \quad (13.9)$$

In the brand-switching example, with limiting state probabilities of  $s_1 = 0.75$  and  $s_2 = 0.25$ , the recurrence times would be

$$\mu_{11} = \frac{1}{s_1} = \frac{1}{0.75} = 1.333$$

$$\mu_{22} = \frac{1}{s_2} = \frac{1}{0.25} = 4.0.$$

Thus, the mean time between purchases of a TCC car would be 1.33 purchases, and the mean time between purchases of a non-TCC car would be 4.0 purchases.

### EXAMPLE 13.3

Consider again the competition between the three Los Angeles TV stations (Example 13.2). Based on the information in that example, the transition probability matrix can be stated as

This Week's Leader	Next Week's Leader		
	A	B	C
A	0.40	0.35	0.25
B	0.10	0.70	0.20
C	0.15	0.25	0.60

- a) If station B is the current ratings leader, how many weeks will it be, on the average, before station A becomes the leader?

We need to calculate  $\mu_{BA}$ . The set of equations to be solved can be formulated, in general terms, according to Equation (13.8).

$$\mu_{BA} = 1 + p_{BB}\mu_{BA} + p_{BC}\mu_{CA}$$

$$\mu_{CA} = 1 + p_{CB}\mu_{BA} + p_{CC}\mu_{CA}$$

Inserting the actual transition probabilities and rewriting,

$$0.30\mu_{BA} - 0.20\mu_{CA} = 1$$

$$-0.25\mu_{BA} + 0.75\mu_{CA} = 1.$$

Multiplying the first equation by 3.75, adding the two equations, and solving,

$$0.875\mu_{BA} = 4.75$$

$$\mu_{BA} = 5.43.$$

On the average, it will be 5.43 weeks until station A takes over the leadership.

br What is the average number of weeks between the time each station is the ratings leader and the very next time it is again the leader?

To answer this question, we must calculate the expected recurrence time for each state. According to Equation (13.9), we need to know only the steady-state probability of being in each possible state. Those values were calculated in Example 13.2 as  $s_A = 0.165$ ,  $s_B = 0.455$ , and  $s_C = 0.347$ . Therefore,

$$\mu_{AA} = \frac{1}{s_A} = \frac{1}{0.165} = 5.95 \text{ weeks.}$$

$$\mu_{BB} = \frac{1}{s_B} = \frac{1}{0.455} = 2.06 \text{ weeks.}$$

$$\mu_{CC} = \frac{1}{s_C} = \frac{1}{0.347} = 2.88 \text{ weeks.}$$

### Systems with Transient and Absorbing States

We will usually be interested in the expected number of steps from a given transient state  $i$  until absorption (by *any* absorbing state) and the probability of reaching a *specific* absorbing state starting from state  $i$ . In the next example (Example 13.4), which involves door-to-door sales, in order to determine the estimated cost per sale, we will need to calculate the expected number of calls until a sale is either completed or lost. We will also have to know the probability of making a successful sale.

Suppose we now combine all absorbing states into a composite absorbing state  $q$ . The first-passage time  $\mu_{iq}$  is then the expected number of steps to absorption.

$$\mu_{iq} = 1 - \sum_{k \neq q} p_{ik}\mu_{kk}. \quad (13.10)$$

where  $q$  represents a composite state consisting of all absorbing states. This equation permits us to form an  $m \times m$  set of linear simultaneous equations, where  $m$  is the number of transient states.

Since we generally need to differentiate between absorbing states, let us calculate the probability of reaching a given absorbing state.

Let  $p'_{ij}$  = probability that absorbing state  $j$  is eventually reached when starting in transient state  $i$ .

Using the same conceptual approach as in the derivation of the first-passage times, and employing the one-step transition probabilities,

$$p'_{ij} = p_{ij} + \sum_{k \neq j} p_{ik} p'_{kj} \quad \text{for all transient states } k. \quad (13.11)$$

The first term is the probability of a direct transition to absorbing state  $j$ . Each product in the summation is the probability of a direct transition to transient state  $k$  multiplied by the probability of eventually getting from state  $k$  to state  $j$ . Again, the result is a set of linear equations, with as many equations and unknowns as there are transient states.

Let us now solve an example problem involving a Markov process with transient and absorbing states.

#### EXAMPLE 13.4

Screem Security Systems markets a standard home security system by using door-to-door salesmen as demonstrators. A sales call will be either to a new prospect or to a prospect previously called on and showing enough interest to warrant another sales call. After the sales call has been made, the result will be either a completed sale, enough interest shown to justify an additional call, or a lost sale. Screem concentrates its sales efforts on likely prospects only.

A salesman receives \$40 for each Screem system sold, and the salesman's estimated cost per call is \$10. We wish to find the expected profit from calling on a new prospect and from calling on an interested prospect.

It will be assumed that the selling process can be represented as a Markov process with the following states:

- State 1: New prospect
- State 2: Interested prospect, based on previous sales calls
- State 3: Sale lost on this call
- State 4: Sale completed on this call

The following probability data are believed to be representative. For a new customer, there is 0.30 probability of losing the sale, 0.20 probability of making the sale, and 0.50 probability that the prospect will be interested enough to make another sales call worthwhile. For a previously interested customer, the probability of losing the sale is 0.25, the probability of completing the sale is 0.50, and the probability of the prospect's retaining interest is 0.25. From these data and assuming that the probabilities applying to an interested customer do not change if he receives more than one sales call before coming to a final decision, the following matrix of one-step transition probabilities can be constructed.

$$P = \begin{bmatrix} & 1 & 2 & 3 & 4 \\ 1 & 0 & 0.50 & 0.30 & 0.20 \\ 2 & 0 & 0.25 & 0.25 & 0.50 \\ 3 & 0 & 0 & 1 & 0 \\ 4 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Since the process always starts in state 1, with a new prospect, the starting row vector is

$$S(0) = [1 \ 0 \ 0 \ 0].$$

To calculate the expected profit, we must determine the expected number of calls until the customer is either sold or lost. We also need to determine the probability that a sale is successfully completed.

Let us start by constructing a transition diagram, Figure 13.5, so as to get a picture of the process. The process always starts in state 1. States 1 and 2 are transient states, since the process eventually ends in either state 3 or state 4, the absorbing states.

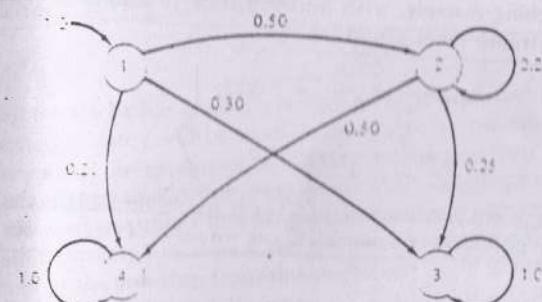


Figure 13.5 Transition Diagram, Screem Security Systems

To find the expected number of calls until the prospect makes a decision, we must combine the two absorbing states into a single composite state. This is done in the transition diagram of Figure 13.6, where the composite absorbing state is denoted as state 5.

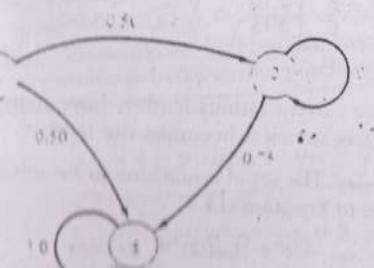


Figure 13.6 Transition Diagram with Composite Absorbing State, Screem Security Systems

## Birth-Death Queueing Systems in Equilibrium

In the previous chapter we studied a variety of stochastic processes. We indicated that Markov processes play a fundamental role in the study of queueing systems, and after presenting the main results from that theory, we then considered a special form of Markov process known as the birth-death process. We also showed that birth-death processes enjoy a most convenient property, namely, that the time between births and the time between deaths (when the system is nonempty) are each exponentially distributed.\* We then developed Eq. (2.127), which gives the basic equations of motion for the general birth-death process with stationary birth and death rates.<sup>†</sup> The solution of this set of equations gives the transient behavior of the queueing process and some important special cases were discussed earlier. In this chapter we study the limiting form of these equations to obtain the equilibrium behavior of birth-death queueing systems.

The importance of elementary queueing theory comes from its historical influence as well as its ability to describe behavior that is to be found in more complex queueing systems. The methods of analysis to be used in this chapter in large part do not carry over to the more involved queueing situations; nevertheless, the obtained results do provide insight into the basic behavior of many of these other queueing systems.

It is necessary to keep in mind how the birth-death process describes queueing systems. As an example, consider a doctor's office made up of a waiting room (in which a queue is allowed to form, unfortunately) and a service facility consisting of the doctor's examination room. Each time a patient enters the waiting room from outside the office we consider this to be an arrival to the queueing system; on the other hand, this arrival may well be considered to be a birth of a new member of a population, where the population consists of all patients present. In a similar fashion, when a patient leaves

\* This comes directly from the fact that they are Markov processes.

<sup>†</sup> In addition to these equations, one requires the conservation relation given in Eq. (2.122) and a set of initial conditions  $\{P_k(0)\}$ .

the office after being treated, he is considered to be a *departure* from the queueing system; in terms of a birth-death process this is considered to be a *death* of a member of the population.

We have considerable freedom in constructing a large number of queueing systems through the choice of the birth coefficients  $\lambda_k$  and death coefficients  $\mu_k$ , as we shall see shortly. First, let us establish the general solution for the equilibrium behavior.

### 3.1. GENERAL EQUILIBRIUM SOLUTION

As we saw in Chapter 2, the time-dependent solution of the birth-death system quickly becomes unmanageable when we consider any sophisticated set of birth-death coefficients. Furthermore, were we always capable of solving for  $P_k(t)$  it is not clear how useful that set of functions would be in aiding our understanding of the behavior of these queueing systems (too much information is sometimes a curse!). Consequently, it is natural for us to ask whether the probabilities  $P_k(t)$  eventually settle down as  $t$  gets large and display no more "transient" behavior. This inquiry on our part is analogous to the questions we asked regarding the existence of  $\pi_k$  in the limit of  $\pi_k(t)$  as  $t \rightarrow \infty$ . For our queueing studies here we choose to denote the limiting probability as  $p_k$  rather than  $\pi_k$ , purely for convenience. Accordingly, let

$$p_k \triangleq \lim_{t \rightarrow \infty} P_k(t) \quad (3.1)$$

where  $p_k$  is interpreted as the limiting probability that the system contains  $k$  members (or equivalently is in state  $E_k$ ) at some arbitrary time in the distant future. The question regarding the existence of these limiting probabilities is of concern to us, but will be deferred at this point until we obtain the general steady-state or limiting solution. It is important to understand that whereas  $p_k$  (assuming it exists) is no longer a function of  $t$ , we are not claiming that the process does not move from state to state in this limiting case; certainly, the number of members in the population will change with time, but the long-run probability of finding the system with  $k$  members will be properly described by  $p_k$ .

Accepting the existence of the limit in Eq. (3.1), we may then set  $\lim dP_k(t)/dt$  as  $t \rightarrow \infty$  equal to zero in the Kolmogorov forward equations (of motion) for the birth-death system [given in Eqs. (2.127)] and immediately obtain the result

$$0 = -(\lambda_k + \mu_k)p_k + \lambda_{k-1}p_{k-1} + \mu_{k+1}p_{k+1} \quad k \geq 1 \quad (3.2)$$

$$0 = -\lambda_0 p_0 + \mu_1 p_1 \quad k = 0 \quad (3.3)$$

The annoying task of providing a separate equation for  $k = 0$  may be overcome by agreeing once and for all that the following birth and death

coefficients are identically equal to 0:

$$\lambda_{-1} = \lambda_{-2} = \lambda_{-3} = \cdots = 0$$

$$\mu_0 = \mu_{-1} = \mu_{-2} = \cdots = 0$$

Furthermore, since it is perfectly clear that we cannot have a negative number of members in our population, we will, in most cases, adopt the convention that

$$p_{-1} = p_{-2} = p_{-3} = \cdots = 0$$

Thus, for all values of  $k$ , we may reformulate Eqs. (3.2) and (3.3) into the following set of difference equations for  $k = \dots, -2, -1, 0, 1, 2, \dots$

$$0 = -(\lambda_k + \mu_k)p_k + \lambda_{k-1}p_{k-1} + \mu_{k+1}p_{k+1} \quad (3.4)$$

We also require the conservation relation

$$\sum_{k=0}^{\infty} p_k = 1 \quad (3.5)$$

Recall from the previous chapter that the limit given in the Eq. (3.1) is independent of the initial conditions.

Just as we used the state-transition-rate diagram as an inspection technique for writing down the equations of motion in Chapter 2, so may we use the same concept in writing down the *equilibrium* equations [Eqs. (3.2) and (3.3)] directly from that diagram. In this equilibrium case it is clear that flow must be *conserved* in the sense that the input flow must equal the output flow from a given state. For example, if we look at Figure 2.9 once again and concentrate on state  $E_k$  in equilibrium, we observe that

$$\text{Flow rate into } E_k = \lambda_{k-1}p_{k-1} + \mu_{k+1}p_{k+1}$$

and

$$\text{Flow rate out of } E_k = (\lambda_k + \mu_k)p_k$$

In equilibrium these two must be the same and so we have immediately

$$\lambda_{k-1}p_{k-1} + \mu_{k+1}p_{k+1} = (\lambda_k + \mu_k)p_k \quad (3.6)$$

But this last is just Eq. (3.4) again! By inspection we have established the equilibrium difference equations for our system. The same comments apply here as applied earlier regarding the conservation of flow across any closed boundary; for example, rather than surrounding each state and writing down its equation we could choose a sequence of boundaries the first of which surrounds  $E_0$ , the second of which surrounds  $E_0$  and  $E_1$ , and so on, each time adding the next higher-numbered state to get a new boundary. In such an example the  $k$ th boundary (which surrounds states  $E_0, E_1, \dots, E_{k-1}$ ) would

2 BIRTH-DEATH QUEUING SYSTEMS IN EQUILIBRIUM

lead to the following simple conservation of flow relationship:

$$\lambda_{k-1}p_{k-1} = \mu_k p_k \quad (3.7)$$

This last set of equations is equivalent to drawing a vertical line separating adjacent states and equating flows across this boundary; this set of difference equations is equivalent to our earlier set.

The solution for  $p_k$  in Eq. (3.4) may be obtained by at least two methods. One way is first to solve for  $p_1$  in terms of  $p_0$  by considering the case  $k = 0$ , that is,

$$p_1 = \frac{\lambda_1}{\mu_1} p_0 \quad (3.8)$$

We may then consider Eq. (3.4) for the case  $k = 1$  and using Eq. (3.8) obtain

$$0 = -(\lambda_1 + \mu_1)p_1 + \lambda_0 p_0 + \mu_2 p_2$$

$$0 = -(\lambda_1 + \mu_1) \frac{\lambda_1}{\mu_1} p_0 + \lambda_0 p_0 + \mu_2 p_2$$

$$0 = -\frac{\lambda_1^2}{\mu_1} p_0 - \lambda_0 p_0 + \lambda_0 p_0 + \mu_2 p_2$$

and so

$$p_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} p_0 \quad (3.9)$$

If we examine Eqs. (3.8) and (3.9) we may justifiably guess that the general solution to Eq. (3.4) must be

$$p_k = \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k} p_0 \quad (3.10)$$

To validate this assertion we need merely use the inductive argument and apply Eq. (3.10) to Eq. (3.4) solving for  $p_{k-1}$ . Carrying out this operation we do, in fact, find that (3.10) is the solution to the general birth-death process in this steady-state or limiting case. We have thus expressed all equilibrium probabilities  $p_k$  in terms of a single unknown constant  $p_0$ :

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \quad k = 0, 1, 2, \dots \quad (3.11)$$

(Recall the usual convention that an empty product is unity by definition.) Equation (3.5) provides the additional condition that allows us to determine  $p_0$ ; thus, summing over all  $k$ , we obtain

$$\sum_{k=0}^{\infty} p_k = p_0 \sum_{k=0}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \quad (3.12)$$

$$\Rightarrow p_0 = \frac{1}{\sum_{k=0}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}}}$$

This "product" solution for  $p_k$  ( $k = 0, 1, 2, \dots$ ) simply obtained, is a principal equation in elementary queueing theory and, in fact, is the point of departure for all of our further solutions in this chapter.

A second easy way to obtain the solution to Eq. (3.4) is to rewrite that equation as follows:

$$\lambda_{k-1}p_{k-1} - \mu_k p_k = \lambda_k p_k - \mu_{k+1}p_{k+1} \quad (3.13)$$

Defining

$$g_k = \lambda_k p_k - \mu_{k+1}p_{k+1} \quad (3.14)$$

we have from Eq. (3.13) that

$$g_{k-1} = g_k \quad \dots \quad (3.15)$$

Clearly Eq. (3.15) implies that

$$g_k = \text{constant with respect to } k \quad (3.16)$$

However, since  $\lambda_{-1} = \mu_0 = 0$ , Eq. (3.14) gives

$$g_{-1} = 0$$

and so the constant in Eq. (3.16) must be 0. Setting  $g_k$  equal to 0, we immediately obtain from Eq. (3.14).

$$p_{k-1} = \frac{\lambda_k}{\mu_{k-1}} p_k \quad (3.17)$$

Solving Eq. (3.17) successively beginning with  $k = 0$  we obtain the earlier solution, namely, Eqs. (3.11) and (3.12).

We now address ourselves to the existence of the steady-state probabilities  $p_k$  given by Eqs. (3.11) and (3.12). Simply stated, in order for those expressions to represent a probability distribution, we usually require that  $p_0 > 0$ . This clearly places a condition upon the birth and death coefficients in those equations. Essentially, what we are requiring is that the system occasionally empties; that this is a condition for stability seems quite reasonable when one interprets it in terms of real life situations.\* More precisely, we may classify the possibilities by first defining the two sums

$$S_1 \triangleq \sum_{k=0}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \quad (3.18)$$

$$S_2 \triangleq \sum_{k=0}^{\infty} \left( 1 / \left( \lambda_k \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_{i+1}} \right) \right) \quad (3.19)$$

\* It is easy to construct counterexamples to this case, and so we require the precise arguments which follow.

All states  $E_k$  of our birth-death process will be ergodic if and only if

$$\text{Ergodic: } S_1 < \infty$$

$$S_2 = \infty$$

On the other hand, all states will be recurrent null if and only if

$$\text{Recurrent null: } S_1 = \infty$$

$$S_2 = \infty$$

Also, all states will be transient if and only if

$$\text{Transient: } S_1 = \infty$$

$$S_2 < \infty$$

It is the ergodic case that gives rise to the equilibrium probabilities  $\{p_k\}$  and that is of most interest to our studies. We note that the condition for ergodicity is met whenever the sequence  $\{\lambda_k/\mu_k\}$  remains below unity from some  $k$  onwards, that is, if there exists some  $k_0$  such that for all  $k \geq k_0$  we have

$$\frac{\lambda_k}{\mu_k} < 1 \quad (3.20)$$

 $\mu_k$ 

We will find this to be true in most of the queueing systems we study.

We are now ready to apply our general solution as given in Eqs. (3.11) and (3.12) to some very important special cases. Before we launch headlong into that discussion, let us put at ease those readers who feel that the birth-death constraints of permitting only nearest-neighbor transitions are too confining. It is true that the solution given in Eqs. (3.11) and (3.12) applies only to nearest-neighbor birth-death processes. However, rest assured that the equilibrium methods we have described can be extended to more general than nearest-neighbor systems; these generalizations are considered in Chapter 4.

### 3.2. M/M/1: THE CLASSICAL QUEUEING SYSTEM

As mentioned in Chapter 2, the celebrated M/M/1 queue is the simplest nontrivial interesting system and may be described by selecting the birth-death coefficients as follows:

$$\lambda_k = \lambda \quad k = 0, 1, 2, \dots$$

$$\mu_k = \mu \quad k = 1, 2, 3, \dots$$

That is, we set all birth\* coefficients equal to a constant  $\lambda$  and all death\*

\* In this case, the average interarrival time is  $\bar{t} = 1/\lambda$  and the average service time is  $\bar{x} = 1/\mu$ ; this follows since both  $\bar{t}$  and  $\bar{x}$  are both exponentially distributed.

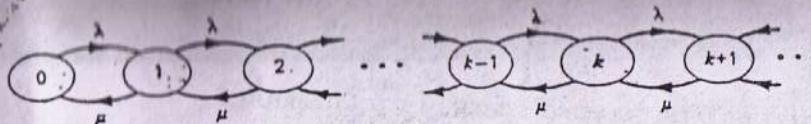


Figure 3.1 State-transition-rate diagram for M/M/1.

coefficients equal to a constant  $\mu$ . We further assume that infinite queueing space is provided and that customers are served in a first-come-first-served fashion (although this last is not necessary for many of our results). For this important example the state-transition-rate diagram is as given in Figure 3.1.

Applying these coefficients to Eq. (3.11) we have

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{\mu}$$

or

$$p_k = p_0 \left( \frac{\lambda}{\mu} \right)^k \quad k \geq 0 \quad (3.21)$$

The result is immediate. The conditions for our system to be ergodic (and, therefore, to have an equilibrium solution  $p_k > 0$ ) are that  $S_1 < \infty$  and  $S_2 = \infty$ ; in this case the first condition becomes

$$S_1 = \sum_{k=0}^{\infty} \frac{\lambda^k}{p_0} = \sum_{k=0}^{\infty} \left( \frac{\lambda}{\mu} \right)^k < \infty$$

The series on the left-hand side of the inequality will converge if and only if  $\lambda/\mu < 1$ . The second condition for ergodicity becomes

$$S_2 = \sum_{k=0}^{\infty} \frac{1}{\lambda(p_k/p_0)} = \sum_{k=0}^{\infty} \frac{1}{\lambda} \left( \frac{\mu}{\lambda} \right)^k = \infty$$

This last condition will be satisfied if  $\lambda/\mu \leq 1$ ; thus the necessary and sufficient condition for ergodicity in the M/M/1 queue is simply  $\lambda < \mu$ . In order to solve for  $p_0$  we use Eq. (3.12) [or Eq. (3.5) as suits the reader] and obtain

$$p_0 = \frac{1}{1 + \sum_{k=1}^{\infty} \left( \frac{\lambda}{\mu} \right)^k}$$

### BIRTH-DEATH QUEUING SYSTEMS IN EQUILIBRIUM

The sum converges since  $\lambda < \mu$  and so

$$p_0 = \frac{1}{1 + \frac{\lambda/\mu}{1 - \lambda/\mu}}$$

$$\boxed{P_0 = 1 - \rho} \quad (3.22)$$

Thus

$$p_k = \frac{\rho^k}{1 - \rho}$$

From Eq. (2.29) we have  $\rho = \lambda/\mu$ . From our stability conditions, we therefore require that  $0 \leq \rho < 1$ ; note that this insures that  $p_0 > 0$ . From Eq. (3.21) we have, finally,

$$p_k = (1 - \rho)\rho^k \quad k = 0, 1, 2, \dots \quad (3.23)$$

Equation (3.23) is indeed the solution for the steady-state probability of finding  $k$  customers in the system.\* We make the important observation that  $p_k$  depends upon  $\lambda$  and  $\mu$  only through their ratio  $\rho$ .

The solution given by Eq. (3.23) for this fundamental system is graphed in Figure 3.2 for the case of  $\rho = 1/2$ . Clearly, this is the geometric distribution (which shares the fundamental memoryless property with the exponential distribution). As we develop the behavior of the M/M/1 queue, we shall continue to see that almost all of its important probability distributions are of the memoryless type.

An important measure of a queuing system is the average number of customers in the system  $\bar{N}$ . This is clearly given by

$$\begin{aligned} \bar{N} &= \sum_{k=0}^{\infty} k p_k \\ &= (1 - \rho) \sum_{k=0}^{\infty} k \rho^k \end{aligned}$$

Using the trick similar to the one used in deriving Eq. (2.142) we have

$$\begin{aligned} \bar{N} &= (1 - \rho) \rho \frac{\frac{d}{d\rho} \sum_{k=0}^{\infty} \rho^k}{\frac{d}{d\rho} \rho} \\ &= (1 - \rho) \rho \frac{1}{1 - \rho} \\ \boxed{\bar{N} = \frac{\rho}{1 - \rho}} \quad (3.24) \end{aligned}$$

\* If we inspect the transient solution for M/M/1 given in Eq. (2.163), we see the term  $(1 - \rho)t\rho^t$ ; the reader may verify that, for  $t < 1$ , the limit of the transient solution agrees with our solution here.

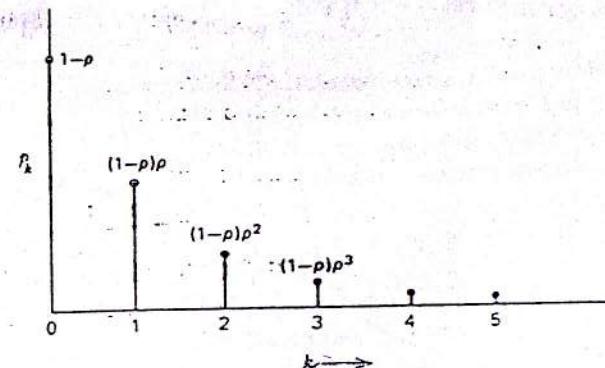


Figure 3.2 The solution for  $p_k$  in the system M/M/1.

The behavior of the expected number in the system is plotted in Figure 3.3. By similar methods we find that the variance of the number in the system is given by

$$\sigma_N^2 = \sum_{k=0}^{\infty} (k - \bar{N})^2 p_k \quad (3.25)$$

$$\sigma_N^2 = \frac{\rho}{(1 - \rho)^2}$$

We may now apply Little's result directly from Eq. (2.25) in order to obtain

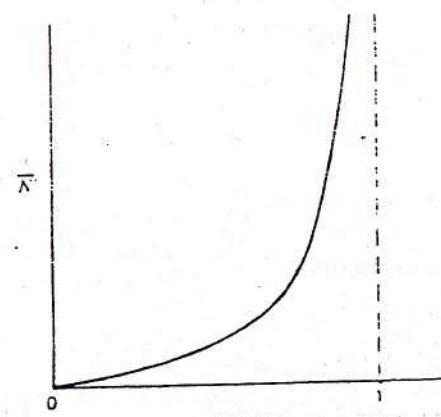


Figure 3.3 The average number in the system M/M/1.

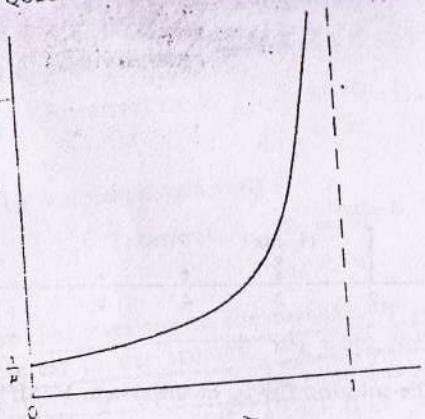


Figure 3.4 Average delay as a function of  $\rho$  for M/M/1.

T, the average time spent in the system as follows:

$$\begin{aligned} T &= \frac{\bar{N}}{\lambda} \\ T &= \left( \frac{\rho}{1 - \rho} \right) \left( \frac{1}{\lambda} \right) \\ T &= \frac{1/\mu}{1 - \rho} \end{aligned} \quad (3.26)$$

This dependence of average time on the utilization factor  $\rho$  is shown in Figure 3.4. The value obtained by  $T$  when  $\rho = 0$  is exactly the average service time expected by a customer; that is, he spends no time in queue and  $1/\mu$  sec in service on the average.

The behavior given by Eqs. (3.24) and (3.26) is rather dramatic. As  $\rho$  approaches unity, both the average number in the system and the average time in the system grow in an unbounded fashion.\* Both these quantities have a

\* We observe at  $\rho = 1$  that the system behavior is unstable; this is not surprising if one recalls that  $\rho < 1$  was our condition for ergodicity. What is perhaps surprising is that the behavior of the average number  $\bar{N}$  and of the average system time  $T$  deteriorates so badly as  $\rho \rightarrow 1$  from below; we had seen for steady flow systems in Chapter 1 that so long as  $R < C$  (which corresponds to the case  $\rho < 1$ ) no queue formed and smooth, rapid flow proceeded through the system. Here in the M/M/1 queue we find this is no longer true and that we pay an extreme penalty when we attempt to run the system near (but below) its capacity. The

### 3.3. DISCOURAGED ARRIVALS

simple pole at  $\rho = 1$ . This type of behavior with respect to  $\rho$  as  $\rho$  approaches 1 is characteristic of almost every queueing system one can encounter. We will see it again in M/G/1 in Chapter 5 as well as in the heavy traffic behavior of G/G/1 (and also in the tight bounds on G/G/1 behavior) in Volume II, Chapter 2.

Another interesting quantity to calculate is the probability of finding at least  $k$  customers in the system:

$$\begin{aligned} P[\geq k \text{ in system}] &= \sum_{i=k}^{\infty} p_i \\ &= \sum_{i=k}^{\infty} (1 - \rho) \rho^i \\ P[\geq k \text{ in system}] &= \rho^k \end{aligned} \quad (3.27)$$

Thus we see that the probability of exceeding some limit on the number of customers in the system is a geometrically decreasing function of that number and decays very rapidly.

With the tools at hand we are now in a position to develop the probability density function for the time spent in the system. However, we defer that development until we treat the more general case of M/G/1 in Chapter 5 [see Eq. (5.118)]. Meanwhile, we proceed to discuss numerous other birth-death queues in equilibrium.

### 3.3. DISCOURAGED ARRIVALS

This next example considers a case where arrivals tend to get discouraged when more and more people are present in the system. One possible way to model this effect is to choose the birth and death coefficients as follows:

$$\begin{aligned} \lambda_k &= \frac{\alpha}{k+1} & k = 0, 1, 2, \dots \\ \mu_k &= \mu & k = 1, 2, 3, \dots \end{aligned}$$

We are here assuming an harmonic discouragement of arrivals with respect to the number present in the system. The state-transition-rate diagram in this

intuitive explanation here is that with random flow (e.g., M/M/1) we get occasional bursts of traffic which temporarily overwhelm the server; while it is still true that the server will be idle on the average  $1 - \rho = p_0$  of the time this average idle time will not be distributed uniformly within small time intervals but will only be true in the long run. On the other hand, in the steady flow case (which corresponds to our system D/D/1) the system idle time will be distributed quite uniformly in the sense that after every service time (of exactly  $1/\mu$  sec) there will be an idle time of exactly  $(1/\lambda - 1/\mu)$  sec. Thus it is the variability in both the interarrival time and in the service time which gives rise to the disastrous behavior near  $\rho = 1$ ; any reduction in the variation of either random variable will lead to a reduction in the average waiting time, as we shall see again and again.

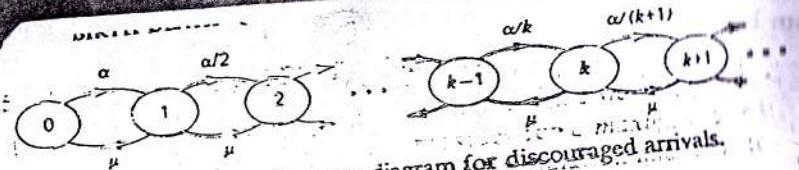


Figure 3.5 State-transition-rate diagram for discouraged arrivals.  
case is as shown in Figure 3.5. We apply Eq. (3.11) immediately to obtain for  $p_k$

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\alpha/(i+1)}{\mu} \quad (3.28)$$

$$p_k = p_0 \left( \frac{\alpha}{\mu} \right)^k \frac{1}{k!} \quad (3.29)$$

Solving for  $p_0$  from Eq. (3.12) we have

$$p_0 = 1 / \left[ 1 + \sum_{k=1}^{\infty} \left( \frac{\alpha}{\mu} \right)^k \frac{1}{k!} \right]$$

$$p_0 = e^{-\alpha/\mu}$$

From Eq. (2.32) we have therefore,

$$p = 1 - e^{-\alpha/\mu} \quad [p = 1 - P] \quad (3.30)$$

Note that the ergodic condition here is merely  $\alpha/\mu < \infty$ . Going back to Eq. (3.29) we have the final solution

$$p_k = \frac{(\alpha/\mu)^k}{k!} e^{-\alpha/\mu} \quad k = 0, 1, 2, \dots \quad (3.31)$$

We thus have a Poisson distribution for the number of customers in the system of discouraged arrivals! From Eqs. (2.131) and (2.132) we have that the expected number in the system is

$$\bar{N} = \sum_{k=0}^{\infty} k p_k = \frac{\alpha}{\mu}$$

In order to calculate  $T$ , the average time spent in the system, we may use Little's result again. For this we require  $\lambda$ , which is directly calculated from  $p = \lambda \bar{x} = \lambda/\mu$ ; thus from Eq. (3.30)

$$\lambda = \mu p = \mu(1 - e^{-\alpha/\mu}) \quad (3.32)$$

Using this\* and Little's result we then obtain

$$T = \frac{\alpha}{\mu^2(1 - e^{-\alpha/\mu})} \quad (3.32)$$

\* Note that this result could have been obtained from  $\lambda = \sum_k \lambda_k p_k$ . The reader should verify this last calculation.

### 3.4. M/M/ $\infty$ : RESPONSIVE SERVERS (INFINITE NUMBER OF SERVERS)

Here we consider the case that may be interpreted either as that of a responsive server who accelerates her service rate linearly when more customers are waiting or may be interpreted as the case where there is always a new clerk or server available for each arriving customer. In particular, we set

$$\begin{aligned} \lambda_k &= \lambda & k = 0, 1, 2, \dots \\ \mu_k &= k\mu & k = 1, 2, 3, \dots \end{aligned}$$

Here the state-transition-rate diagram is that shown in Figure 3.6. Going directly to Eq. (3.11) for the solution we obtain

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu} \quad (3.33)$$

Need we go any further? The reader should compare Eq. (3.33) with Eq. (3.28). These two are in fact equivalent, and so we immediately have the solutions for  $p_k$  and  $\bar{N}$ ,

$$p_k = \frac{(\lambda/\mu)^k}{k!} e^{-\lambda/\mu} \quad k = 0, 1, 2, \dots \quad (3.34)$$

$$\bar{N} = \frac{\lambda}{\mu}$$

Here, too, the ergodic condition is simply  $\lambda/\mu < \infty$ . It appears then that a system of discouraged arrivals behaves exactly the same as a system that includes a responsive server. However, Little's result provides a different (and simpler) form for  $T$  here than that given in Eq. (3.32): thus

$$T = \frac{1}{\mu}$$

This answer is, of course, obvious since if we use the interpretation where each arriving customer is granted his own server, then his time in system will be merely his service time which clearly equals  $1/\mu$  on the average.

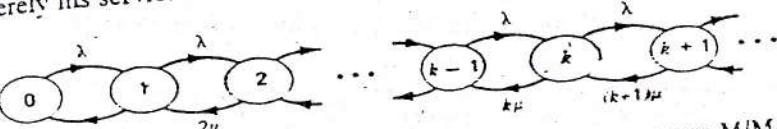


Figure 3.6 State-transition-rate diagram for the infinite-server case M/M/ $\infty$ .

Komal

### 3.5 M/M/m: THE m-SERVER CASE

Here again we consider a system with an unlimited waiting room and with a constant arrival rate  $\lambda$ . The system provides for a maximum of  $m$  servers. This is within the reach of our birth-death formulation and leads to

$$\lambda_k = \lambda \quad k = 0, 1, 2, \dots$$

$$\mu_k = \min [k\mu, m\mu] \\ = \begin{cases} k\mu & 0 \leq k \leq m \\ m\mu & m \leq k \end{cases}$$

From Eq. (3.20) it is easily seen that the condition for ergodicity is  $\lambda/m\mu < 1$ . The state-transition-rate diagram is shown in Figure 3.7. When we go to solve for  $p_k$  from Eq. (3.11) we find that we must separate the solution into two parts, since the dependence of  $\mu_k$  upon  $k$  is also in two parts. Accordingly, for  $k \leq m$ ,

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu} \\ = p_0 \left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!} \quad (3.35)$$

Similarly, for  $k \geq m$ ,

$$p_k = p_0 \prod_{i=0}^{m-1} \frac{\lambda}{(i+1)\mu} \prod_{i=m}^{k-1} \frac{\lambda}{m\mu} \\ = p_0 \left(\frac{\lambda}{\mu}\right)^k \frac{1}{m! m^{k-m}} \quad (3.36)$$

Collecting together the results from Eqs. (3.35) and (3.36) we have

$$p_k = \begin{cases} p_0 \frac{(mp)^k}{k!} & k \leq m \\ p_0 \frac{(\rho)^k m^m}{m!} & k \geq m \end{cases} \quad (3.37)$$

where

$$\rho = \frac{\lambda}{m\mu} < 1 \quad (3.38)$$

This expression for  $\rho$  follows that in Eq. (2.30) and is consistent with our

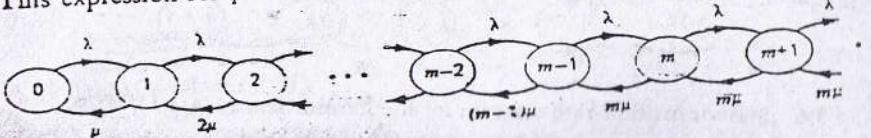


Figure 3.7 State-transition-rate diagram for M/M/m.

definition in terms of the expected fraction of busy servers. We may now solve for  $p_0$  from Eq. (3.12), which gives us

$$p_0 = \left[ 1 + \sum_{k=1}^{m-1} \frac{(mp)^k}{k!} + \sum_{k=m}^{\infty} \frac{(mp)^k}{m!} \frac{1}{m^{k-m}} \right]^{-1}$$

and so

$$p_0 = \left[ \sum_{k=0}^{m-1} \frac{(mp)^k}{k!} + \left( \frac{(mp)^m}{m!} \right) \left( \frac{1}{1-\rho} \right) \right]^{-1} \quad (3.39)$$

The probability that an arriving customer is forced to join the queue is given by

$$P[\text{queueing}] = \sum_{k=m}^{\infty} p_k \\ = \sum_{k=m}^{\infty} p_0 \frac{(mp)^k}{m!} \frac{1}{m^{k-m}}$$

Thus

$$P[\text{queueing}] = \frac{\left( \frac{(mp)^m}{m!} \right) \left( \frac{1}{1-\rho} \right)}{\left[ \sum_{k=0}^{m-1} \frac{(mp)^k}{k!} + \left( \frac{(mp)^m}{m!} \right) \left( \frac{1}{1-\rho} \right) \right]} \quad (3.40)$$

This probability is of wide use in telephony and gives the probability that no trunk (i.e., server) is available for an arriving call (customer) in a system of  $m$  trunks; it is referred to as *Erlang's C formula* and is often denoted\* by  $C(m, \lambda/\mu)$ .

### 3.6 M/M/1/K: FINITE STORAGE

We now consider for the first time the case of a queuing system in which there is a maximum number of customers that may be stored: in particular, we assume the system can hold at most a total of  $K$  customers (including the customer in service) and that any further arriving customers will in fact be refused entry to the system and will depart immediately without service. Newly arriving customers will continue to be generated according to a Poisson process but only those who find the system with strictly less than  $K$  customers will be allowed entry. In telephony the refused customers are considered to be "lost"; for the system in which  $K = 1$  (i.e., no waiting room at all) this is referred to as a "blocked calls cleared" system with a single server.

\* Europeans use the symbol  $E_{k,m}(\lambda/\mu)$ .

### BIRTH-DEATH QUEUING SYSTEMS IN EQUILIBRIUM

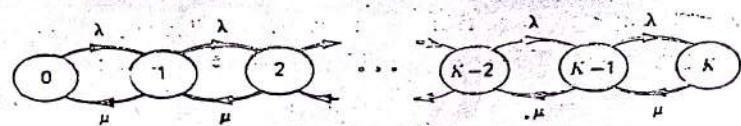


Figure 3.8 State-transition-rate diagram for the case of finite storage room  $M/M/1/K$ .

It is interesting that we are capable of accommodating this seemingly complex system description with our birth-death model. In particular, we accomplish this by effectively "turning off" the Poisson input as soon as the system fills up, as follows:

$$\begin{aligned}\lambda_k &= \begin{cases} \lambda & k < K \\ 0 & k \geq K \end{cases} \\ \mu_k &= \mu \quad k = 1, 2, \dots, K\end{aligned}$$

From Eq. (3.20), we see that this system is always ergodic. The state-transition-rate diagram for this finite Markov chain is shown in Figure 3.8. Proceeding directly with Eq. (3.11) we obtain

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{\mu} \quad k \leq K$$

or

$$p_k = p_0 \left( \frac{\lambda}{\mu} \right)^k \quad k \leq K \quad (3.41)$$

Of course, we also have

$$p_k = 0 \quad k > K \quad (3.42)$$

In order to solve for  $p_0$  we use Eqs. (3.41) and (3.42) in Eq. (3.12) to obtain

$$\begin{aligned}p_0 &= \left[ 1 + \sum_{k=1}^K \left( \frac{\lambda}{\mu} \right)^k \right]^{-1} \\ &= \left[ 1 + \frac{(\lambda/\mu)(1 - (\lambda/\mu)^K)}{1 - \lambda/\mu} \right]^{-1}\end{aligned}$$

and so

$$p_0 = \frac{1 - \lambda/\mu}{1 - (\lambda/\mu)^{K+1}}$$

Thus, finally,

$$p_k = \begin{cases} \frac{1 - \lambda/\mu}{1 - (\lambda/\mu)^{K+1}} \left( \frac{\lambda}{\mu} \right)^k & 0 \leq k \leq K \\ 0 & \text{otherwise} \end{cases} \quad (3.43)$$

### 3.7. M/M/m/m: m-SERVER LOSS SYSTEMS

105



Figure 3.9 State-transition-rate diagram for  $m$ -server loss system  $M/M/m/m$ .

For the case of blocked calls cleared ( $K = 1$ ) we have

$$p_k = \begin{cases} \frac{1}{1 + \lambda/\mu} & k = 0 \\ \frac{\lambda/\mu}{1 + \lambda/\mu} & k = 1 = K \\ 0 & \text{otherwise} \end{cases} \quad (3.44)$$

### 3.7. M/M/m/m: m-SERVER LOSS SYSTEMS

Here we have again a blocked calls cleared situation in which there are available  $m$  servers. Each newly arriving customer is given his private server; however, if a customer arrives when all servers are occupied, that customer is lost. We create this artifact as above by choosing the following birth and death coefficients:

$$\begin{aligned}\lambda_k &= \begin{cases} \lambda & k < m \\ 0 & k \geq m \end{cases} \\ \mu_k &= k\mu \quad k = 1, 2, \dots, m\end{aligned}$$

Here again, ergodicity is always assured. This finite state-transition-rate diagram is shown in Figure 3.9.

Applying Eq. (3.11) we obtain

$$\boxed{p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda}{(i+1)\mu} \quad k \leq m}$$

or

$$p_k = \begin{cases} p_0 \left( \frac{\lambda}{\mu} \right)^k \frac{1}{k!} & k \leq m \\ 0 & k > m \end{cases} \quad (3.45)$$

Solving for  $p_0$  we have

$$p_0 = \left[ \sum_{k=0}^m \left( \frac{\lambda}{\mu} \right)^k \frac{1}{k!} \right]^{-1}$$

This particular system is of great interest to those in telephony [so much so that a special case of Eq. (3.45) has been tabulated and graphed in many books

on telephony]. Specifically,  $p_m$  describes the fraction of time that all  $m$  servers are busy. The name given to this probability expression is Erlang's loss formula and it is given by

$$p_m = \frac{(\lambda/\mu)^m / m!}{\sum_{k=0}^m (\lambda/\mu)^k / k!} \quad (3.46)$$

This equation is also referred to as Erlang's B formula and is commonly denoted\* by  $B(m, \lambda/\mu)$ . Formula (3.46) was first derived by Erlang in 1917!

### 3.8. M/M/1//M†: FINITE CUSTOMER POPULATION— SINGLE SERVER

Here we consider the case where we no longer have a Poisson input process with an infinite user population, but rather have a finite population of possible users. The system structure is such that we have a total of  $M$  users: a customer is either in the system (consisting of a queue and a single server) or outside the system and in some sense "arriving." In particular, when a customer is in the "arriving" condition then the time it takes him to arrive is a random variable with an exponential distribution whose mean is  $1/\lambda$  sec. All customers act independently of each other. As a result, when there are  $k$  customers in the system (queue plus service) then there are  $M - k$  customers in the arriving state and, therefore, the total average arrival rate in this state is  $\lambda(M - k)$ . We see that this system is in a strong sense self-regulating. By this we mean that when the system gets busy, with many of these customers in the queue, then the rate at which additional customers arrive is in fact reduced, thus lowering the further congestion of the system. We model this quite appropriately with our birth-death process choosing for parameters

$$\lambda_k = \begin{cases} \lambda(M - k) & 0 \leq k \leq M \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_k = \mu \quad k = 1, 2, \dots$$

The system is ergodic. We assume that we have sufficient room to contain  $M$  customers in the system. The finite state-transition-rate diagram is shown in Figure 3.10. Using Eq. (3.11) we solve for  $p_k$  as follows:

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda(M - i)}{\mu} \quad 0 \leq k \leq M$$

\* Europeans use the notation  $E_{1,m}(\lambda/\mu)$ .

† Recall that a blank entry in either of the last two optional positions in this notation means an entry of  $\infty$ ; thus here we have the system M/M/1// $\infty$ /M.

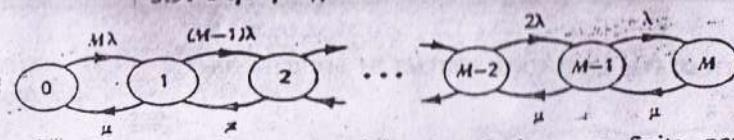


Figure 3.10 State-transition-rate diagram for single-server finite population system M/M/1//M.

Thus

$$p_k = \begin{cases} p_0 \left( \frac{\lambda}{\mu} \right)^k \frac{M!}{(M - k)!} & 0 \leq k \leq M \\ 0 & k > M \end{cases} \quad (3.47)$$

In addition, we obtain for  $p_0$

$$p_0 = \left[ \sum_{k=0}^M \left( \frac{\lambda}{\mu} \right)^k \frac{M!}{(M - k)!} \right]^{-1} \quad (3.48)$$

### 3.9. M/M/ $\infty$ //M: FINITE CUSTOMER POPULATION— "INFINITE" NUMBER OF SERVERS

We again consider the finite population case, but now provide a separate server for each customer in the system. We model this as follows:

$$\lambda_k = \begin{cases} \lambda(M - k) & 0 \leq k \leq M \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_k = k\mu \quad k = 1, 2, \dots$$

Clearly, this too is an ergodic system. The finite state-transition-rate diagram is shown in Figure 3.11. Solving this system, we have from Eq. (3.11)

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda(M - i)}{(i + 1)\mu} = p_0 \left( \frac{\lambda}{\mu} \right)^k \binom{M}{k} \quad 0 \leq k \leq M \quad (3.49)$$

where the binomial coefficient is defined in the usual way,

$$\binom{M}{k} \triangleq \frac{M!}{k!(M - k)!}$$

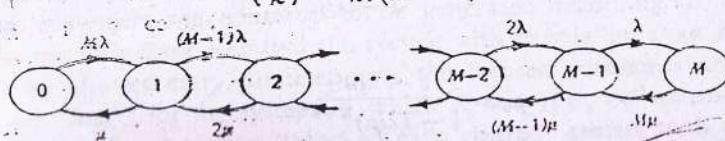


Figure 3.11 State-transition-rate diagram for "infinite"-server finite population system M/M/∞//M.

Solving for  $p_0$  we have

$$p_0 = \left[ \sum_{k=0}^M \left( \frac{\lambda}{\mu} \right)^k \binom{M}{k} \right]^{-1}$$

and so

$$p_0 = \frac{1}{(1 + \lambda/\mu)^M}$$

Thus

$$p_k = \begin{cases} \left( \frac{\lambda}{\mu} \right)^k \binom{M}{k} & 0 \leq k \leq M \\ \frac{1}{(1 + \lambda/\mu)^M} & \text{otherwise} \end{cases} \quad (3.50)$$

We may easily calculate the expected number of people in the system from

$$\begin{aligned} \bar{N} &= \sum_{k=0}^M k p_k \\ &= \frac{\sum_{k=0}^M k \left( \frac{\lambda}{\mu} \right)^k \binom{M}{k}}{(1 + \lambda/\mu)^M} \end{aligned}$$

Using the partial-differentiation trick such as for then have

$$\bar{N} = \frac{M\lambda/\mu}{1 + \lambda/\mu}$$

### 3.10. M/M/m/K/M: FINITE POPULATION, m-SERVER CASE, FINITE STORAGE

This rather general system is the most complicated we have so far considered and will reduce to all of the previous cases (except the example of discouraged arrivals) as we permit the parameters of this system to vary. We assume we have a finite population of  $M$  customers, each with an "arriving" parameter  $\lambda$ . In addition, the system has  $m$  servers, each with parameter  $\mu$ . The system also has finite storage room such that the total number of customers in the system (queueing plus those in service) is no more than  $K$ . We assume  $M \geq K \geq m$ ; customers arriving to find  $K$  already in the system are "lost" and return immediately to the arriving state as if they had just completed service. This leads to the following set of birth-death coefficients:

$$\lambda_k = \begin{cases} \lambda(M - k) & 0 \leq k \leq K - 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_k = \begin{cases} k\mu & 0 \leq k \leq m \\ m\mu & k \geq m \end{cases}$$

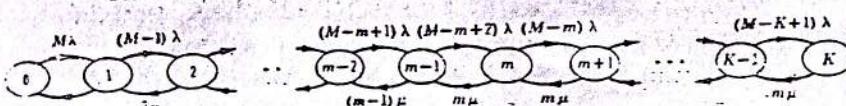


Figure 3.12 State-transition-rate diagram for  $m$ -server, finite storage, finite population system  $M/M/m/K/M$ .

In Figure 3.12 we see the most complicated of our finite state-transition-rate diagrams. In order to apply Eq. (3.11) we must consider two regions. First, for the range  $0 \leq k \leq m - 1$  we have

$$\begin{aligned} p_k &= p_0 \prod_{i=0}^{k-1} \frac{\lambda(M-i)}{(i+1)\mu} \\ &= p_0 \left( \frac{\lambda}{\mu} \right)^k \binom{M}{k} \quad 0 \leq k \leq m-1 \end{aligned} \quad (3.51)$$

For the region  $m \leq k \leq K$  we have

$$\begin{aligned} p_k &= p_0 \prod_{i=0}^{m-1} \frac{\lambda(M-i)}{(i+1)\mu} \prod_{i=m}^{k-1} \frac{\lambda(M-i)}{m\mu} \\ &= p_0 \left( \frac{\lambda}{\mu} \right)^k \binom{M}{k} \frac{k!}{m!} m^{m-k} \quad m \leq k \leq K \end{aligned} \quad (3.52)$$

The expression for  $p_0$  is rather complex and will not be given here, although it may be computed in a straightforward manner. In the case of a pure loss system (i.e.,  $M \geq K = m$ ), the stationary state probabilities are given by

$$p_k = \frac{\binom{M}{k} \left( \frac{\lambda}{\mu} \right)^k}{\sum_{i=0}^m \binom{M}{i} \left( \frac{\lambda}{\mu} \right)^i} \quad k = 0, 1, \dots, m \quad (3.53)$$

This is known as the Engset distribution.

We could continue these examples ad nauseam but we will instead take a benevolent approach and terminate the set of examples here. Additional examples are given in the exercises. It should be clear to the reader by now that a large number of interesting queueing structures can be modeled with the birth-death process. In particular, we have demonstrated the ability to model the multiple-server case, the finite-population case, the finite-storage case and combinations thereof. The common element in all of these is that the solution for the equilibrium probabilities  $\{p_k\}$  is given in Eqs. (3.11) and (3.12). Only systems whose solutions are given by these equations have been considered in this chapter. However, there are many other Markovian systems that lend themselves to simple solution and which are important in queueing

theory. In the next chapter (4) we consider the equilibrium solution for Markovian queues; in Chapter 5 we will generalize to semi-Markov processes in which the service time distribution  $B(x)$  is permitted to be general, and in Chapter 6 we revert back to the exponential service time case, but permit the interarrival time distribution  $A(t)$  to be general; in both of these cases an imbedded Markov chain will be identified and solved. Only when both  $A(t)$  and  $B(x)$  are nonexponential do we require the methods of advanced queueing theory discussed in Chapter 8. (There are some special nonexponential distributions that may be described with the theory of Markov processes, and these too are discussed in Chapter 4.)

### EXERCISES

- 3.1. Consider a pure Markovian queueing system in which

$$\begin{aligned}\lambda_k &= \begin{cases} \lambda & 0 \leq k \leq K \\ 2\lambda & K < k \end{cases} \\ \mu_k &= u \quad k = 1, 2, \dots\end{aligned}$$

- (a) Find the equilibrium probabilities  $p_k$  for the number in the system.
- (b) What relationship must exist among the parameters of the problem in order that the system be stable and, therefore, that this equilibrium solution in fact be reached? Interpret this answer in terms of the possible dynamics of the system.

- 3.2. Consider a Markovian queueing system in which

$$\begin{aligned}\lambda_k &= x^k \lambda \quad k \geq 0, 0 \leq x < 1 \\ \mu_k &= u \quad k \geq 1\end{aligned}$$

- (a) Find the equilibrium probability  $p_k$  of having  $k$  customers in the system. Express your answer in terms of  $p_0$ .
- (b) Give an expression for  $p_0$ .

- 3.3. Consider an M/M/2 queueing system where the average arrival rate is  $\lambda$  customers per second and the average service time is  $1/\mu$  sec, where  $\lambda < 2\mu$ .

- (a) Find the differential equations that govern the time-dependent probabilities  $P_k(t)$ .
- (b) Find the equilibrium probabilities

$$p_k = \lim_{t \rightarrow \infty} P_k(t)$$

- 3.4. Consider an M/M/1 system with parameters  $\lambda, \mu$  in which customers are impatient. Specifically, upon arrival, customers estimate their queueing time  $w$  and then join the queue with probability  $e^{-w}$  (or leave with probability  $1 - e^{-w}$ ). The estimate is  $w = k/\mu$  when the new arrival finds  $k$  in the system. Assume  $0 \leq \alpha$ .

- (a) In terms of  $p_0$ , find the equilibrium probabilities  $p_k$  of finding  $k$  in the system. Give an expression for  $p_0$  in terms of the system parameters.
- (b) For  $0 < x, 0 < \mu$  under what conditions will the equilibrium solution hold?
- (c) For  $x \rightarrow \infty$ , find  $p_0$  explicitly and find the average number in the system.

- 3.5. Consider a birth-death system with the following birth and death coefficients:

$$\begin{aligned}\lambda_k &= (k - 2)\lambda & k = 0, 1, 2, \dots \\ \mu_k &= k\mu & k = 1, 2, 3, \dots\end{aligned}$$

All other coefficients are zero.

- (a) Solve for  $p_k$ . Be sure to express your answer explicitly in terms of  $\lambda, \mu$ , and  $u$  only.
- (b) Find the average number of customers in the system.

- 3.6. Consider a birth-death process with the following coefficients:

$$\begin{aligned}\lambda_k &= \alpha k(K_2 - k) & k = K_1, K_1 + 1, \dots, K_2 \\ \mu_k &= \beta k(k - K_1) & k = K_1, K_1 + 1, \dots, K_2\end{aligned}$$

where  $K_1 \leq K_2$  and where these coefficients are zero outside the range  $K_1 \leq k \leq K_2$ . Solve for  $p_k$  (assuming that the system initially contains  $K_1 \leq k \leq K_2$  customers).

- 3.7. Consider an M/M/m system that is to serve the pooled sum of two Poisson arrival streams; the  $i$ th stream has an average arrival rate given by  $\lambda_i$  and exponentially distributed service times with mean  $1/\mu_i$  ( $i = 1, 2$ ). The first stream is an ordinary stream whereby each arrival requires exactly one of the  $m$  servers; if all  $m$  servers are busy then any newly arriving customer of type 1 is lost. Customers from the second class each require the simultaneous use of  $m_0$  servers (and will occupy them all simultaneously for the same exponentially distributed amount of time whose mean is  $1/\mu_2$  sec); if a customer from this class finds less than  $m_0$  idle servers then he too is lost to the system. Find the fraction of type 1 customers and the fraction of type 2 customers that are lost.