Quantitative Module **D**

Waiting-Line Models

Module Outline

LEARNING OBJECTIVES

When you complete this module you should be able to

IDENTIFY OR DEFINE:

The assumptions of the four basic waiting-line models

DESCRIBE OR EXPLAIN:

How to apply waiting-line models

How to conduct an economic analysis of queues

Paris's EuroDisney, Tokyo's Disney Japan, and the U.S.'s Disney World and Disneyland all have one feature in common—long lines and seemingly endless waits. However, Disney is one of the world's leading companies in the scientific analysis of queuing theory. It analyzes queuing behaviors and can predict which rides will draw what length crowds. To keep visitors happy, Disney makes lines appear to be constantly moving forward, entertains people while they wait, and posts signs telling visitors how many minutes until they reach each ride.

Queuing theory
A body of knowledge about waiting lines.

Waiting line (queue)
Items or people in a line awaiting service.

The body of knowledge about waiting lines, often called **queuing theory**, is an important part of operations and a valuable tool for the operations manager. **Waiting lines** are a common situation—they may, for example, take the form of cars waiting for repair at a Midas Muffler Shop, copying jobs waiting to be completed at a Kinko's print shop, or vacationers waiting to enter Mr. Toad's Wild Ride at Disney. Table D.1 lists just a few OM uses of waiting-line models.

Waiting-line models are useful in both manufacturing and service areas. Analysis of queues in terms of waiting-line length, average waiting time, and other factors helps us to understand service systems (such as bank teller stations), maintenance activities (that might repair broken machinery), and shop-floor control activities. Indeed, patients waiting in a doctor's office and broken drill presses waiting in a repair facility have a lot in common from an OM perspective. Both use human and equipment resources to restore valuable production assets (people and machines) to good condition.

TABLE D.1 ■
Common Queuing Situations

SITUATION	ARRIVALS IN QUEUE	SERVICE PROCESS
Supermarket	Grocery shoppers	Checkout clerks at cash register
Highway toll booth	Automobiles	Collection of tolls at booth
Doctor's office	Patients	Treatment by doctors and nurses
Computer system	Programs to be run	Computer processes jobs
Telephone company	Callers	Switching equipment forwards calls
Bank	Customers	Transactions handled by teller
Machine maintenance	Broken machines	Repair people fix machines
Harbor	Ships and barges	Dock workers load and unload

CHARACTERISTICS OF A WAITING-LINE SYSTEM

In this section, we take a look at the three parts of a waiting-line, or queuing, system (as shown in Figure D.1):

1. *Arrivals or inputs to the system.* These have characteristics such as population size, behavior, and a statistical distribution.
2. *Queue discipline, or the waiting line itself.* Characteristics of the queue include whether it is limited or unlimited in length and the discipline of people or items in it.
3. *The service facility.* Its characteristics include its design and the statistical distribution of service times.

We now examine each of these three parts.

Arrival Characteristics

The input source that generates arrivals or customers for a service system has three major characteristics:

1. *Size* of the arrival population.
2. *Behavior* of arrivals.
3. *Pattern* of arrivals (statistical distribution).

Unlimited, or **infinite, population**
A queue in which a virtually unlimited number of people or items could request the services, or in which the number of customers or arrivals on hand at any given moment is a very small portion of potential arrivals.

Size of the Arrival (Source) Population Population sizes are considered either unlimited (essentially infinite) or limited (finite). When the number of customers or arrivals on hand at any given moment is just a small portion of all potential arrivals, the arrival population is considered **unlimited**, or **infinite**. Examples of unlimited populations include cars arriving at a big-city carwash, shoppers arriving at a supermarket, and students arriving to register for classes at a large university. Most queuing models assume such an infinite arrival population. An example of a **limited**, or **finite**, population is found in a copying shop that has, say, eight copying machines. Each of the copiers is a potential "customer" that may break down and require service.

Limited, or **finite, population**
A queue in which there are only a limited number of potential users of the service.

Pattern of Arrivals at the System Customers arrive at a service facility either according to some known schedule (for example, one patient every 15 minutes or one student every half hour) or else they arrive *randomly*. Arrivals are considered random when they are independent of one another and their occurrence cannot be predicted exactly. Frequently in queuing problems, the number of arrivals per unit of time can be estimated by a probability distribution known as the

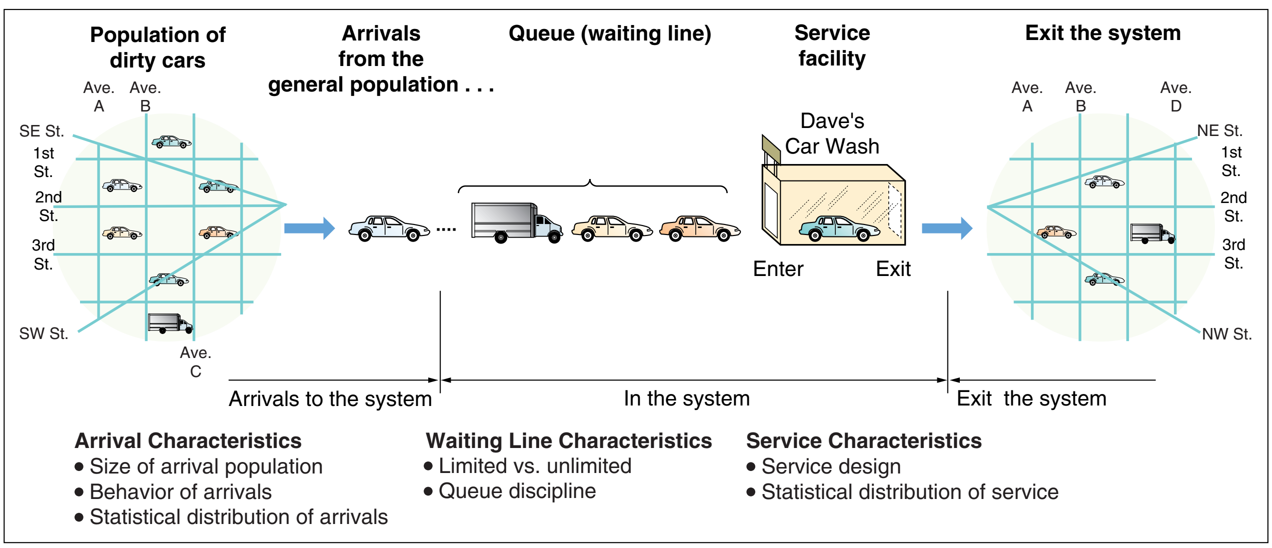


FIGURE D.1 ■ Three Parts of a Waiting Line, or Queuing System, at Dave's Car Wash

Poisson distribution
A discrete probability distribution that often describes the arrival rate in queuing theory.

Poisson distribution.[1] For any given arrival time (such as 2 customers per hour or 4 trucks per minute), a discrete Poisson distribution can be established by using the formula

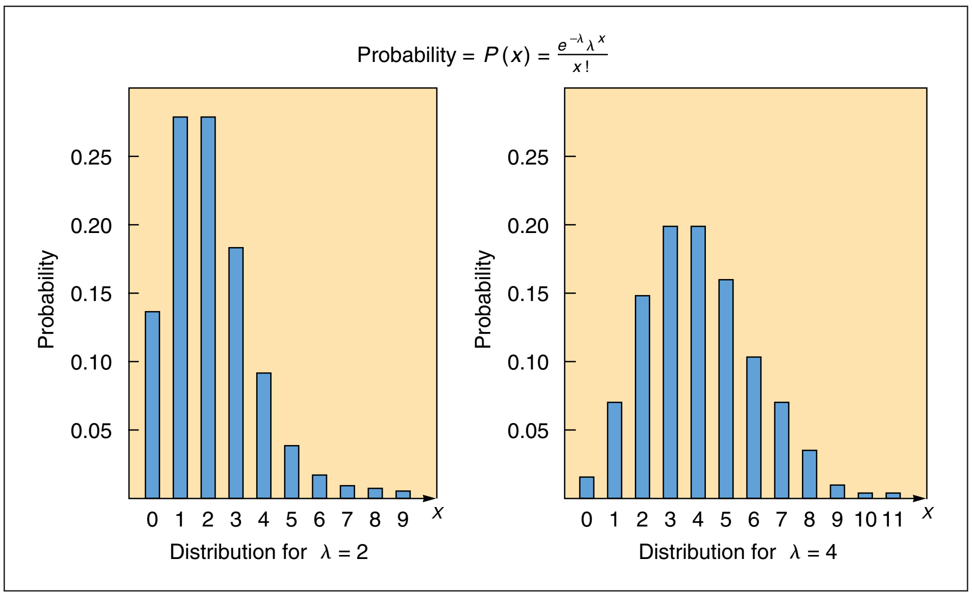
$$P(x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad \text{for } x = 0, 1, 2, 3, 4, \ldots \tag{D-1}$$

where $P(x)$ = probability of x arrivals
x = number of arrivals per unit of time
λ = average arrival rate
e = 2.7183 (which is the base of the natural logarithms)

With the help of the table in Appendix II, which gives the value of $e^{-\lambda}$ for use in the Poisson distribution, these values are easy to compute. Figure D.2 illustrates the Poisson distribution for $\lambda = 2$ and $\lambda = 4$. This means that if the average arrival rate is $\lambda = 2$ customers per hour, the probability of 0 customers arriving in any random hour is about 13%, probability of 1 customer is about 27%, 2 customers about 27%, 3 customers about 18%, 4 customers about 9%, and so on. The chances that 9 or more will arrive are virtually nil. Arrivals, of course, are not always Poisson distributed (they may follow some other distribution). Patterns, therefore, should be examined to make certain that they are well approximated by Poisson before that distribution is applied.

FIGURE D.2 ■

Two Examples of the Poisson Distribution for Arrival Times



"The other line always moves faster."
Etorre's Observation

"If you change lines, the one you just left will start to move faster than the one you are now in."
O'Brien's Variation

Behavior of Arrivals Most queuing models assume that an arriving customer is a patient customer. Patient customers are people or machines that wait in the queue until they are served and do not switch between lines. Unfortunately, life is complicated by the fact that people have been known to balk or to renege. Customers who *balk* refuse to join the waiting line because it is too long to suit their needs or interests. *Reneging* customers are those who enter the queue but then become impatient and leave without completing their transaction. Actually, both of these situations just serve to highlight the need for queuing theory and waiting-line analysis.

Waiting-Line Characteristics

The waiting line itself is the second component of a queuing system. The length of a line can be either limited or unlimited. A queue is *limited* when it cannot, either by law or because of physical restrictions, increase to an infinite length. A small barbershop, for example, will have only a limited

[1]When the arrival rates follow a Poisson process with mean arrival rate λ, the time between arrivals follows a negative exponential distribution with mean time between arrivals of $1/\lambda$. The negative exponential distribution, then, is also representative of a Poisson process but describes the time between arrivals and specifies that these time intervals are completely random.

number of waiting chairs. Queuing models are treated in this module under an assumption of *unlimited* queue length. A queue is *unlimited* when its size is unrestricted, as in the case of the toll booth serving arriving automobiles.

A second waiting-line characteristic deals with *queue discipline*. This refers to the rule by which customers in the line are to receive service. Most systems use a queue discipline known as the **first-in, first-out (FIFO) rule**. In a hospital emergency room or an express checkout line at a supermarket, however, various assigned priorities may preempt FIFO. Patients who are critically injured will move ahead in treatment priority over patients with broken fingers or noses. Shoppers with fewer than 10 items may be allowed to enter the express checkout queue (but are *then* treated as first-come, first-served). Computer-programming runs also operate under priority scheduling. In most large companies, when computer-produced paychecks are due on a specific date, the payroll program gets highest priority.[2]

First-in, first-out (FIFO) rule
A queuing discipline in which the first customers in line receive the first service.

Service Characteristics

The third part of any queuing system are the service characteristics. Two basic properties are important: (1) design of the service system and (2) the distribution of service times.

Basic Queuing System Designs Service systems are usually classified in terms of their number of channels (for example, number of servers) and number of phases (for example, number of service stops that must be made). A **single-channel queuing system**, with one server, is typified by the drive-in bank with only one open teller. If, on the other hand, the bank has several tellers on duty, with each customer waiting in one common line for the first available teller, then we would have a **multiple-channel queuing system**. Most banks today are multichannel service systems, as are most large barbershops, airline ticket counters, and post offices.

Single-channel queuing system
A service system with one line and one server.

Multiple-channel queuing system
A service system with one waiting line but with several servers.

In a **single-phase system**, the customer receives service from only one station and then exits the system. A fast-food restaurant in which the person who takes your order also brings your food and takes your money is a single-phase system. So is a driver's license agency in which the person taking your application also grades your test and collects your license fee. However, say the restaurant requires you to place your order at one station, pay at a second, and pick up your food at a third. In this case, it is a **multiphase system**. Likewise, if the driver's license agency is large or busy, you will probably have to wait in one line to complete your application (the first service stop), queue again to have your test graded, and finally go to a third counter to pay your fee. To help you relate the concepts of channels and phases, Figure D.3 presents four possible channel configurations.

Single-phase system
A system in which the customer receives service from only one station and then exits the system.

Multiphase system
A system in which the customer receives services from several stations before exiting the system.

Service Time Distribution Service patterns are like arrival patterns in that they may be either constant or random. If service time is constant, it takes the same amount of time to take care of each customer. This is the case in a machine-performed service operation such as an automatic car wash. More often, service times are randomly distributed. In many cases, we can assume that random service times are described by the **negative exponential probability distribution**.

Negative exponential probability distribution
A continuous probability distribution often used to describe the service time in a queuing system.

Figure D.4 shows that if *service times* follow a negative exponential distribution, the probability of any very long service time is low. For example, when an average service time is 20 minutes (or three customers per hour), seldom if ever will a customer require more than 1.5 hours in the service facility. If the mean service time is 1 hour, the probability of spending more than 3 hours in service is quite low.

Measuring the Queue's Performance

Queuing models help managers make decisions that balance service costs with waiting-line costs. Queuing analysis can obtain many measures of a waiting-line system's performance, including the following:

1. Average time that each customer or object spends in the queue.
2. Average queue length.
3. Average time that each customer spends in the system (waiting time plus service time).
4. Average number of customers in the system.
5. Probability that the service facility will be idle.
6. Utilization factor for the system.
7. Probability of a specific number of customers in the system.

[2]The term *FIFS* (first-in, first-served) is often used in place of FIFO. Another discipline, LIFS (last-in, first-served) also called last-in, first-out (LIFO), is common when material is stacked or piled so that the items on top are used first.

FIGURE D.3 ■ Basic Queuing System Designs

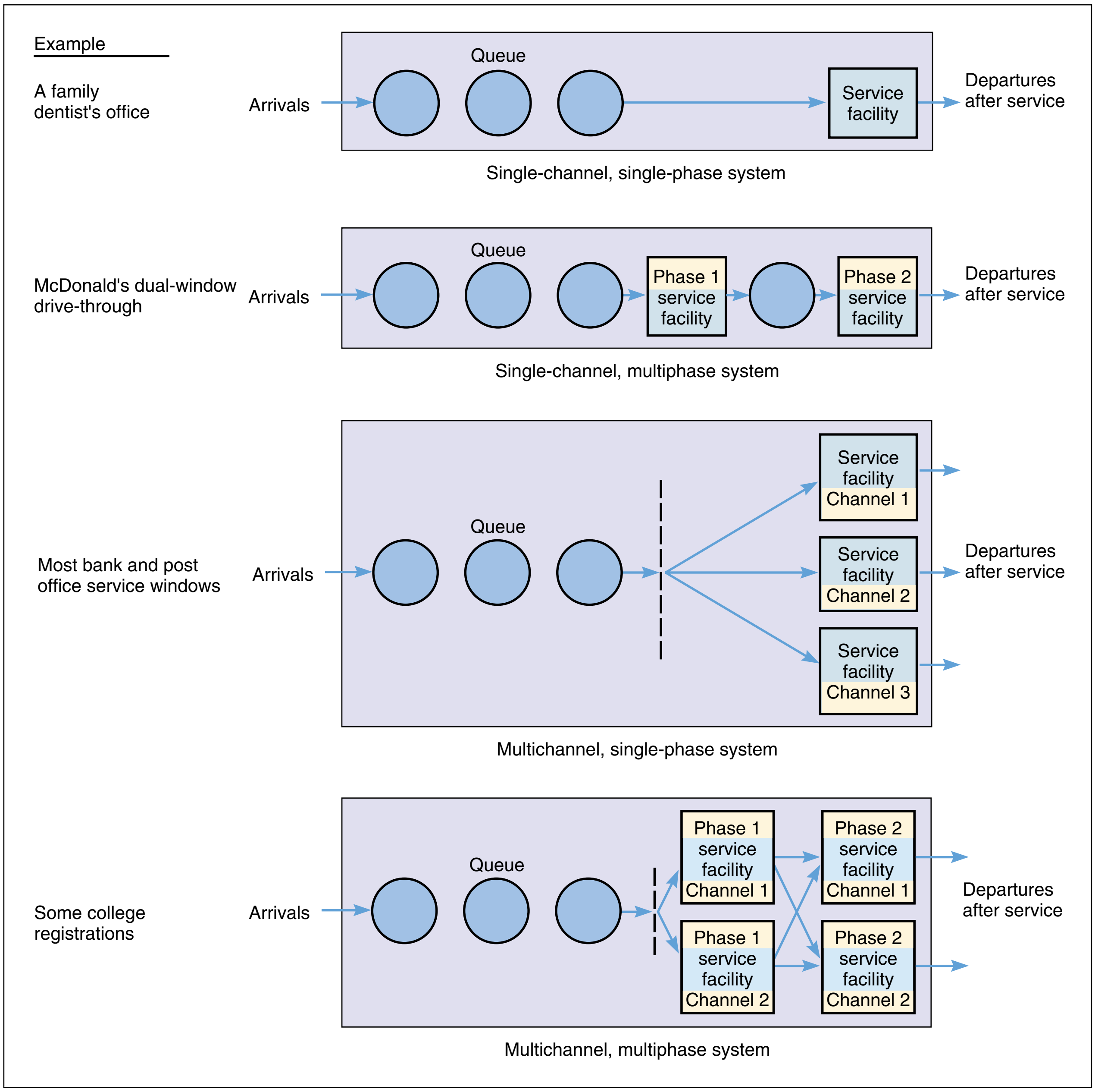
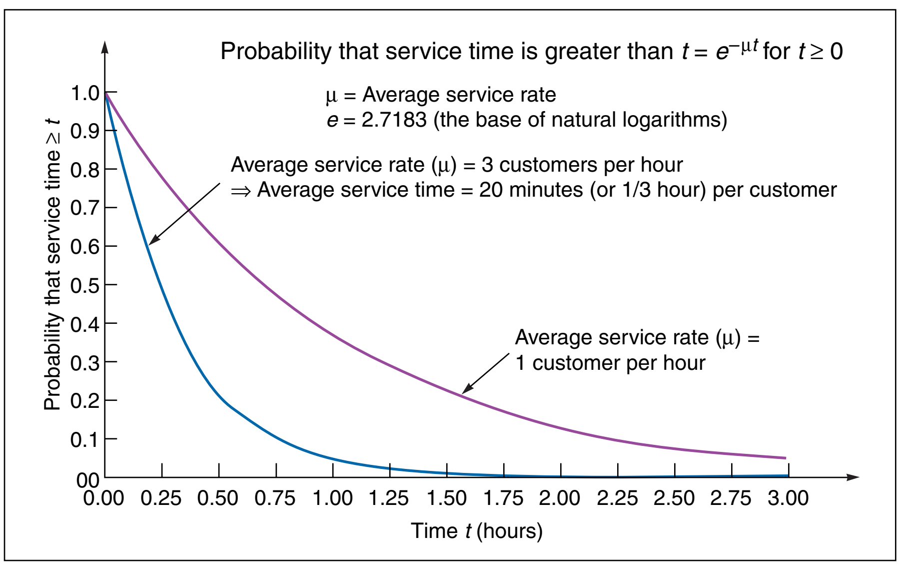


FIGURE D.4 ■

Two Examples of the Negative Exponential Distribution for Service Times

Although Poisson and exponential distributions are commonly used to describe arrival rates and service times, normal and Erlang distributions, or others, may be more valid in certain cases.



OM IN ACTION

L.L. Bean Turns to Queuing Theory

L.L. Bean faced severe problems. It was the peak selling season, and the service level to incoming calls was simply unacceptable. Widely known as a high-quality outdoor goods retailer, about 65% of L.L. Bean's sales volume is generated through telephone orders via its toll-free service centers located in Maine.

Here is how bad the situation was: During certain periods, 80% of the calls received a busy signal, and those who did not often had to wait up to 10 minutes before speaking with a sales agent. L.L. Bean estimated it lost $10 million in profit because of the way it allocated telemarketing resources. Keeping customers waiting "in line" (on the phone) was costing $25,000 per day. On exceptionally busy days, the total orders lost because of queuing problems approached $500,000 in gross revenues.

Developing queuing models similar to those presented here, L.L. Bean was able to set the number of phone lines and the number of agents to have on duty for each half hour of every day of the season. Within a year, use of the model resulted in 24% more calls answered, 17% more orders taken, and 16% more revenues. The new system also meant 81% fewer abandoned callers and an 84% faster answering time. The percent of callers spending less than 20 seconds in the queue increased from 25% to 77%. Needless to say, queuing theory changed the way L.L. Bean thought about telecommunications.

Sources: Modern Material Handling (December 1997): S12–S14; and *Interfaces* (January/February 1991): 75–91 and (March/April 1993): 14–20.

QUEUING COSTS

As described in the *OM in Action* box "L.L. Bean Turns to Queuing Theory," operations managers must recognize the trade-off that takes place between two costs: the cost of providing good service and the cost of customer or machine waiting time. Managers want queues that are short enough so that customers do not become unhappy and either leave without buying or buy but never return. However, managers may be willing to allow some waiting if it is balanced by a significant savings in service costs.

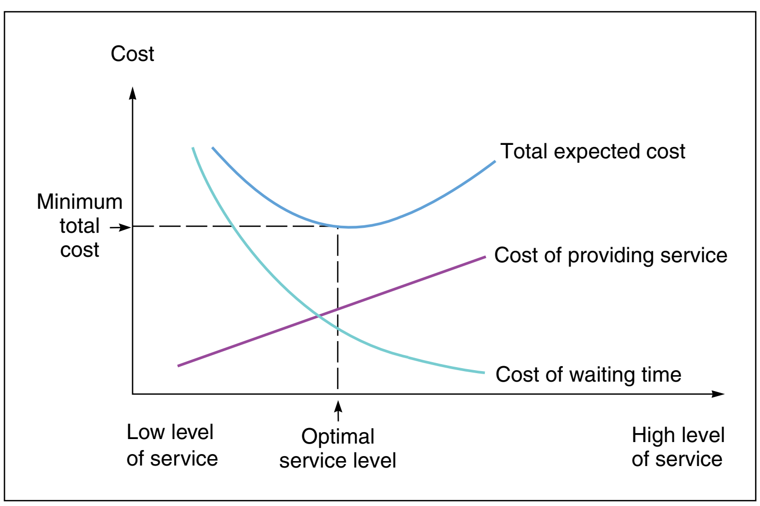
One means of evaluating a service facility is to look at total expected cost. Total cost is the sum of expected service costs plus expected waiting costs.

What does the long wait in the typical doctor's office tell you about the doctor's perception of your cost of waiting?

As you can see in Figure D.5, service costs increase as a firm attempts to raise its level of service. Managers in *some* service centers can vary capacity by having standby personnel and machines that they can assign to specific service stations to prevent or shorten excessively long lines. In grocery stores, for example, managers and stock clerks can open extra checkout counters. In banks and airport check-in points, part-time workers may be called in to help. As the level of service improves (that is, speeds up), however, the cost of time spent waiting in lines decreases. (Refer again to Figure D.5.) Waiting cost may reflect lost productivity of workers while tools or machines await repairs or may simply be an estimate of the cost of customers lost because of poor service and long queues. In some service systems (for example, an emergency ambulance service), the cost of long waiting lines may be intolerably high.

FIGURE D.5 ■

The Trade-Off between Waiting Costs and Service Costs



Different organizations place different values on their customers' time, don't they?

TABLE D.2 ■ Queuing Models Described in This Chapter

MODEL	NAME (TECHNICAL NAME IN PARENTHESES)	EXAMPLE	NUMBER OF CHANNELS	NUMBER OF PHASES	ARRIVAL RATE PATTERN	SERVICE TIME PATTERN	POPULATION SIZE	QUEUE DISCIPLINE
A	Single-channel system (M/M/1)	Information counter at department store	Single	Single	Poisson	Exponential	Unlimited	FIFO
B	Multichannel (M/M/S)	Airline ticket counter	Multi-channel	Single	Poisson	Exponential	Unlimited	FIFO
C	Constant service (M/D/1)	Automated car wash	Single	Single	Poisson	Constant	Unlimited	FIFO
D	Limited population (finite population)	Shop with only a dozen machines that might break	Single	Single	Poisson	Exponential	Limited	FIFO

THE VARIETY OF QUEUING MODELS

Visit a bank or a drive-through restaurant and time arrivals to see what kind of distribution (Poisson or other) they might reflect.

A wide variety of queuing models may be applied in operations management. We will introduce you to four of the most widely used models. These are outlined in Table D.2, and examples of each follow in the next few sections. More complex models are described in queuing theory textbooks[3] or can be developed through the use of simulation (the topic of module F). Note that all four queuing models listed in Table D.2 have three characteristics in common. They all assume

1. Poisson distribution arrivals.
2. FIFO discipline.
3. A single-service phase.

In addition, they all describe service systems that operate under steady, ongoing conditions. This means that arrival and service rates remain stable during the analysis.

Model A (M/M/1): Single-Channel Queuing Model with Poisson Arrivals and Exponential Service Times

The most common case of queuing problems involves the *single-channel*, or single-server, waiting line. In this situation, arrivals form a single line to be serviced by a single station (see Figure D.3 on p. 748). We assume that the following conditions exist in this type of system:

1. Arrivals are served on a first-in, first-out (FIFO) basis, and every arrival waits to be served, regardless of the length of the line or queue.
2. Arrivals are independent of preceding arrivals, but the average number of arrivals (*arrival rate*) does not change over time.
3. Arrivals are described by a Poisson probability distribution and come from an infinite (or very, very large) population.
4. Service times vary from one customer to the next and are independent of one another, but their average rate is known.
5. Service times occur according to the negative exponential probability distribution.
6. The service rate is faster than the arrival rate.

What is the impact of equal service and arrival rates?

When these conditions are met, the series of equations shown in Table D.3 can be developed. Examples D1 and D2 illustrate how Model A (which in technical journals is known as the M/M/1 model) may be used.[4]

[3]See, for example, N. U. Prabhu, *Foundations of Queuing Theory*, Kluwer Academic Publishers (1997).

[4]In queuing notation, the first letter refers to the arrivals (where M stands for Poisson distribution); the second letter refers to service (where M is again a Poisson distribution, which is the same as an exponential rate for service—and a D is a constant service rate); the third symbol refers to the number of servers. So an M/D/1 system (our Model C) has Poisson arrivals, constant service, and one server.

TABLE D.3 ■ Queuing Formulas for Model A: Single-Channel System, Also Called M/M/1

λ = mean number of arrivals per time period

μ = mean number of people or items served per time period

L_s = average number of units (customers) in the system (waiting and being served)

$$= \frac{\lambda}{\mu - \lambda}$$

W_s = average time a unit spends in the system (waiting time plus service time)

$$= \frac{1}{\mu - \lambda}$$

L_q = average number of units waiting in the queue

$$= \frac{\lambda^2}{\mu(\mu - \lambda)}$$

W_q = average time a unit spends waiting in the queue

$$= \frac{\lambda}{\mu(\mu - \lambda)}$$

ρ = utilization factor for the system

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

P_0 = probability of 0 units in the system (that is, the service unit is idle)

$$= 1 - \frac{\lambda}{\mu}$$

$P_{n>k}$ = probability of more than k units in the system, where n is the number of units in the system

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

Example D1

A single-channel queue

Excel OM Data File ModDExD1.xla

Active Model D.1

Example D1 is further illustrated in Active Model D.1 on your CD-ROM.

Tom Jones, the mechanic at Golden Muffler Shop, is able to install new mufflers at an average rate of 3 per hour (or about 1 every 20 minutes), according to a negative exponential distribution. Customers seeking this service arrive at the shop on the average of 2 per hour, following a Poisson distribution. They are served on a first-in, first-out basis and come from a very large (almost infinite) population of possible buyers.

From this description, we are able to obtain the operating characteristics of Golden Muffler's queuing system:

$$\lambda = 2 \text{ cars arriving per hour}$$

$$\mu = 3 \text{ cars serviced per hour}$$

$$L_s = \frac{\lambda}{\mu - \lambda} = \frac{2}{3-2} = \frac{2}{1}$$

$$= 2 \text{ cars in the system, on average}$$

$$W_s = \frac{1}{\mu - \lambda} = \frac{1}{3-2} = 1$$

$$= 1\text{-hour average waiting time in the system}$$

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{2^2}{3(3-2)} = \frac{4}{3(1)} = \frac{4}{3}$$

$$= 1.33 \text{ cars waiting in line, on average}$$

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{2}{3(3-2)} = \frac{2}{3} \text{ hour}$$

$$= 40\text{-minute average waiting time per car}$$

$$\rho = \frac{\lambda}{\mu} = \frac{2}{3}$$

$$= 66.6\% \text{ of time mechanic is busy}$$

$$P_0 = 1 - \frac{\lambda}{\mu} = 1 - \frac{2}{3}$$

$$= .33 \text{ probability there are 0 cars in the system}$$

Probability of More Than *k* Cars in the System

k	$P_{n>k} = (2/3)^{k+1}$
0	.667 ← Note that this is equal to $1 - P_0 = 1 - .33 = .667$.
1	.444
2	.296
3	.198 ← Implies that there is a 19.8% chance that more than 3 cars are in the system.
4	.132
5	.088
6	.058
7	.039

Once we have computed the operating characteristics of a queuing system, it is often important to do an economic analysis of their impact. Although the waiting-line model described above is valuable in predicting potential waiting times, queue lengths, idle times, and so on, it does not identify optimal decisions or consider cost factors. As we saw earlier, the solution to a queuing problem may require management to make a trade-off between the increased cost of providing better service and the decreased waiting costs derived from providing that service.

Example D2 examines the costs involved in Example D1.

Example D2

Economic analysis of example D1

The owner of the Golden Muffler Shop estimates that the cost of customer waiting time, in terms of customer dissatisfaction and lost goodwill, is \$10 per hour of time spent *waiting* in line. Because the average car has a $\frac{2}{3}$-hour wait (W_q) and because there are approximately 16 cars serviced per day (2 arrivals per hour times 8 working hours per day), the total number of hours that customers spend waiting each day for mufflers to be installed is

$$\frac{2}{3}(16) = \frac{32}{3} = 10\frac{2}{3} \text{ hour}$$

L_q and W_q are the two most important queuing parameters when it comes to actual cost analysis.

Hence, in this case,

$$\text{Customer waiting-time cost} = \$10\left(10\frac{2}{3}\right) = \$106.67 \text{ per day}$$

The only other major cost that Golden's owner can identify in the queuing situation is the salary of Jones, the mechanic, who earns \$7 per hour, or \$56 per day. Thus:

$$\begin{aligned}\text{Total expected costs} &= \$106.67 + \$56 \\ &= \$162.67 \text{ per day}\end{aligned}$$

This approach will be useful in Solved Problem D.2 on page 761.

A $P_{n>3}$ of .0625 means that the chance of having more than 3 customers in an airport check-in line at a certain time of day is 1 in 16. If this British Airways office can live with 4 or more passengers in line about 6% of the time, one service agent will suffice. If not, more check-in positions and staff will have to be added.

Model B (M/M/S): Multiple-Channel Queuing Model

Now let's turn to a multiple-channel queuing system in which two or more servers or channels are available to handle arriving customers. We still assume that customers awaiting service form one single line and then proceed to the first available server. Multichannel, single-phase waiting lines are found in many banks today: A common line is formed, and the customer at the head of the line proceeds to the first free teller. (Refer to Figure D.3 on p. 748 for a typical multichannel configuration.)

The multiple-channel system presented in Example D3 again assumes that arrivals follow a Poisson probability distribution and that service times are exponentially distributed. Service is first-come, first-served, and all servers are assumed to perform at the same rate. Other assumptions listed earlier for the single-channel model also apply.

The queuing equations for Model B (which also has the technical name M/M/S) are shown in Table D.4. These equations are obviously more complex than those used in the single-channel model; yet they are used in exactly the same fashion and provide the same type of information as the simpler model. (*Note:* The POM for Windows and Excel OM software described later in this chapter can prove very useful in solving multiple-channel, as well as other, queuing problems.)

TABLE D.4 ■

Queuing Formulas for Model B: Multichannel System, Also Called M/M/S

M = number of channels open
λ = average arrival rate
μ = average service rate at each channel

The probability that there are zero people or units in the system is

$$P_0 = \frac{1}{\left[\sum_{n=0}^{M-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n\right] + \frac{1}{M!}\left(\frac{\lambda}{\mu}\right)^M \frac{M\mu}{M\mu - \lambda}} \quad \text{for } M\mu > \lambda$$

The average number of people or units in the system is

$$L_s = \frac{\lambda\mu(\lambda/\mu)^M}{(M-1)!(M\mu - \lambda)^2} P_0 + \frac{\lambda}{\mu}$$

The average time a unit spends in the waiting line and being serviced (namely, in the system) is

$$W_s = \frac{\mu(\lambda/\mu)^M}{(M-1)!(M\mu - \lambda)^2} P_0 + \frac{1}{\mu} = \frac{L_s}{\lambda}$$

The average number of people or units in line waiting for service is

$$L_q = L_s - \frac{\lambda}{\mu}$$

The average time a person or unit spends in the queue waiting for service is

$$W_q = W_s - \frac{1}{\mu} = \frac{L_q}{\lambda}$$

Example D3

A multiple-channel queue

Excel OM Data File ModDExD3.xla

The Golden Muffler Shop has decided to open a second garage bay and hire a second mechanic to handle installations. Customers, who arrive at the rate of about $\lambda = 2$ per hour, will wait in a single line until 1 of the 2 mechanics is free. Each mechanic installs mufflers at the rate of about $\mu = 3$ per hour.

To find out how this system compares with the old single-channel waiting-line system, we will compute several operating characteristics for the $M = 2$ channel system and compare the results with those found in Example D1:

$$P_0 = \frac{1}{\left[\sum_{n=0}^{1} \frac{1}{n!}\left(\frac{2}{3}\right)^n\right] + \frac{1}{2!}\left(\frac{2}{3}\right)^2 \frac{2(3)}{2(3) - 2}}$$

$$= \frac{1}{1 + \frac{2}{3} + \frac{1}{2}\left(\frac{4}{9}\right)\left(\frac{6}{6-2}\right)} = \frac{1}{1 + \frac{2}{3} + \frac{1}{3}} = \frac{1}{2}$$

$= .5$ probability of zero cars in the system

Active Model D.2

Examples D2 and D3 are further illustrated in Active Model D.2 on the CD-ROM and in the Exercise on page 763.

Then,

$$L_s = \frac{(2)(3)(2/3)^2}{1!\,[2(3)-2]^2}\left(\frac{1}{2}\right)+\frac{2}{3} = \frac{8/3}{16}\left(\frac{1}{2}\right)+\frac{2}{3} = \frac{3}{4}$$

= .75 average number of cars in the system

$$W_s = \frac{L_s}{\lambda} = \frac{3/4}{2} = \frac{3}{8}\text{ hour}$$

= 22.5 minutes average time a car spends in the system

$$L_q = L_s - \frac{\lambda}{\mu} = \frac{3}{4}-\frac{2}{3} = \frac{9}{12}-\frac{8}{12} = \frac{1}{12}$$

= .083 average number of cars in the queue (waiting)

$$W_q = \frac{L_q}{\lambda} = \frac{.083}{2} = .0415\text{ hour}$$

= 2.5 minutes average time a car spends in the queue (waiting)

We can summarize the characteristics of the 2-channel model in Example D3 and compare them to those of the single-channel model in Example D1 as follows:

	SINGLE CHANNEL	TWO CHANNELS
P_0	.33	.5
L_s	2 cars	.75 car
W_s	60 minutes	22.5 minutes
L_q	1.33 cars	.083 car
W_q	40 minutes	2.5 minutes

The increased service has a dramatic effect on almost all characteristics. For instance, note that the time spent waiting in line drops from 40 minutes to only 2.5 minutes.

Use of Waiting Line Tables Imagine the work a manager would face in dealing with $M = 3$, 4, or 5 channel waiting line models if a computer was not readily available. The arithmetic becomes increasingly troublesome. Fortunately, much of the burden of manually examining multiple channel queues can be avoided by using Table D.5. This table, the result of hundreds of computations, represents the relationship between three things: (1) service facility utilization factor, ρ (which is simple to find—it's just λ/μ), (2) number of service channels open, and (3) the average number of customers in the queue, L_q (which is what we'd like to find). For any combination of utilization rate (ρ) and $M = 1, 2, 3, 4$, or 5 open service channels, you can quickly look in the body of the table to read off the appropriate value for L_q.

Example D4 illustrates the use of Table D.5.

Example D4

Use of waiting line tables

Alaska National Bank is trying to decide how many drive-in teller windows to open on a busy Saturday. CEO Ted Eschenbach estimates that customers arrive at a rate of about $\lambda = 18$ per hour, and that each teller can service about $\mu = 20$ customers per hour. Then the utilization rate is $\rho = \lambda/\mu = {}^{18}\!/\!_{20} = .90$. Turning to Table D.5, under $\rho = .90$, Ted sees that if only $M = 1$ service window is open, the average number of customers in line will be 8.1. If two windows are open, L_q drops to .2285 customers, to .03 for $M = 3$ tellers, and to .0041 for $M = 4$ tellers. Adding more open windows at this point will result in an average queue length of 0.

It is also a simple matter to compute the average waiting time in the queue, W_q, since $W_q = L_q/\lambda$. When one channel is open, W_q = 8.1 customers/(18 customers per hour) = .45 hours = 27 minutes waiting time; when two tellers are open, W_q = .2285 customers/(18 customers per hour) = .0127 hours $\cong$ ¾ minute; and so on.

You might also wish to check the calculations in Example D3 against tabled values just to practice the use of Table D.5. You may need to interpolate if your exact ρ value is not found in the first column. Other common operating characteristics besides L_q are published in tabled form in queuing theory textbooks.

TABLE D.5

Values of L_q for $M = 1$–5 Service Channels and Selected Values of $\rho = \lambda/\mu$

POISSON ARRIVALS, EXPONENTIAL SERVICE TIMES

	NUMBER OF SERVICE CHANNELS, *M*				
ρ	1	2	3	4	5
.10	.0111				
.15	.0264	.0008			
.20	.0500	.0020			
.25	.0833	.0039			
.30	.1285	.0069			
.35	.1884	.0110			
.40	.2666	.0166			
.45	.3681	.0239	.0019		
.50	.5000	.0333	.0030		
.55	.6722	.0449	.0043		
.60	.9000	.0593	.0061		
.65	1.2071	.0767	.0084		
.70	1.6333	.0976	.0112		
.75	2.2500	.1227	.0147		
.80	3.2000	.1523	.0189		
.85	4.8166	.1873	.0239	.0031	
.90	8.1000	.2285	.0300	.0041	
.95	18.0500	.2767	.0371	.0053	
1.0		.3333	.0454	.0067	
1.2		.6748	.0904	.0158	
1.4		1.3449	.1778	.0324	.0059
1.6		2.8444	.3128	.0604	.0121
1.8		7.6734	.5320	.1051	.0227
2.0			.8888	.1739	.0398
2.2			1.4907	.2770	.0659
2.4			2.1261	.4305	.1047
2.6			4.9322	.6581	.1609
2.8			12.2724	1.0000	.2411
3.0				1.5282	.3541
3.2				2.3856	.5128
3.4				3.9060	.7365
3.6				7.0893	1.0550
3.8				16.9366	1.5184
4.0					2.2164
4.2					3.3269
4.4					5.2675
4.6					9.2885
4.8					21.6384

Queues exist not only in every industry but also around the world. Here, the Moscow McDonald's on Pushkin Square, four blocks from the Kremlin, boasts 700 indoor and 200 outdoor seats, employs 800 Russian citizens, and generates annual revenues of $80 million. In spite of its size and volume, it still has queues and has had to develop a strategy for dealing with them.

Model C (M/D/1): Constant-Service-Time Model

Some service systems have constant, instead of exponentially distributed, service times. When customers or equipment are processed according to a fixed cycle, as in the case of an automatic car wash or an amusement park ride, constant service times are appropriate. Because constant rates are certain, the values for L_q, W_q, L_s, and W_s are always less than they would be in Model A, which has variable service rates. As a matter of fact, both the average queue length and the average waiting time in the queue are halved with Model C. Constant-service-model formulas are given in Table D.6. Model C also has the technical name M/D/1 in the literature of queuing theory.

TABLE D.6 ■

Queuing Formulas for Model C: Constant Service, Also Called M/D/1

Average length of queue: $L_q = \dfrac{\lambda^2}{2\mu(\mu - \lambda)}$

Average waiting time in queue: $W_q = \dfrac{\lambda}{2\mu(\mu - \lambda)}$

Average number of customers in system: $L_s = L_q + \dfrac{\lambda}{\mu}$

Average waiting time in system: $W_s = W_q + \dfrac{1}{\mu}$

Example D5 gives a constant-service-time analysis.

Example D5

A constant-service model

Excel OM Data File ModDExD5.xla

Active Model D.3

Example D5 is further illustrated in Active Model D.3 on your CD-ROM.

Garcia-Golding Recycling, Inc., collects and compacts aluminum cans and glass bottles in New York City. Its truck drivers currently wait an average of 15 minutes before emptying their loads for recycling. The cost of driver and truck time while they are in queues is valued at \$60 per hour. A new automated compactor can be purchased to process truckloads at a constant rate of 12 trucks per hour (that is, 5 minutes per truck). Trucks arrive according to a Poisson distribution at an average rate of 8 per hour. If the new compactor is put in use, the cost will be amortized at a rate of \$3 per truck unloaded. The firm hires a summer college intern, who conducts the following analysis to evaluate the costs versus benefits of the purchase:

Current waiting cost/trip = (1/4 hr waiting now) (\$60/hr cost) = \$15/trip

New system: $\lambda = 8$ trucks/hr. arriving $\mu = 12$ trucks/hr served

Average waiting time in queue $= W_q = \dfrac{\lambda}{2\mu(\mu - \lambda)} = \dfrac{8}{2(12)(12 - 8)} = \dfrac{1}{12}$ hr

Waiting cost/trip with new compactor = (1/12 hr wait) (\$60/hr cost) = \$ 5/trip

Savings with new equipment = \$15 (current system) – \$5 (new system) = \$10/trip

Cost of new equipment amortized: = \$ 3/trip

Net savings: \$ 7/trip

Model D: Limited-Population Model

When there is a limited population of potential customers for a service facility, we must consider a different queuing model. This model would be used, for example, if we were considering equipment repairs in a factory that has 5 machines, if we were in charge of maintenance for a fleet of 10 commuter airplanes, or if we ran a hospital ward that has 20 beds. The limited-population model allows any number of repair people (servers) to be considered.

This model differs from the three earlier queuing models because there is now a *dependent* relationship between the length of the queue and the arrival rate. Let's illustrate the extreme situation: If your factory had five machines and all were broken and awaiting repair, the arrival rate would drop to zero. In general, then, as the *waiting line* becomes longer in the limited population model, the *arrival rate* of customers or machines drops.

Table D.7 displays the queuing formulas for the limited-population model. Note that they employ a different notation than Models A, B, and C. To simplify what can become time-consuming calculations, finite queuing tables have been developed that determine D and F. D represents the probability that a machine needing repair will have to wait in line. F is a waiting-time efficiency factor. D and F are needed to compute most of the other finite model formulas.

TABLE D.7 ■

Queuing Formulas and Notation for Model D: Limited Population Formulas

Service factor: $X = \dfrac{T}{T+U}$	Average number running: $J = NF(1 - X)$
Average number waiting: $L = N(1 - F)$	Average number being serviced: $H = FNX$
Average waiting time: $W = \dfrac{L(T+U)}{N-L} = \dfrac{T(1-F)}{XF}$	Number of population: $N = J + L + H$

NOTATION

D = probability that a unit will have to wait in queue
F = efficiency factor
H = average number of units being served
J = average number of units not in queue or in service bay
L = average number of units waiting for service
M = number of service channels
N = number of potential customers
T = average service time
U = average time between unit service requirements
W = average time a unit waits in line
X = service factor

Source: L. G. Peck and R. N. Hazelwood, *Finite Queuing Tables* (New York: John Wiley, 1958).

A small part of the published finite queuing tables is illustrated in this section. Table D.8 provides data for a population of $N = 5$.[5]

To use Table D.8, we follow four steps:

1. Compute X (the service factor, where $X = T/(T + U)$.
2. Find the value of X in the table and then find the line for M (where M is the number of service channels).
3. Note the corresponding values for D and F.
4. Compute L, W, J, H, or whichever are needed to measure the service system's performance.

Example D6 illustrates these steps.

Example D6

A limited-population model

Excel OM Data File ModDExD6.xla

Past records indicate that each of the 5 laser computer printers at the U.S. Department of Energy, in Washington, DC, needs repair after about 20 hours of use. Breakdowns have been determined to be Poisson distributed. The one technician on duty can service a printer in an average of 2 hours, following an exponential distribution. Printer downtime costs \$120 per hour. Technicians are paid \$25 per hour. Should the DOE hire a second technician?

Assuming the second technician can repair a printer in an average of 2 hours, we can use Table D.8 (because there are $N = 5$ machines in this limited population) to compare the costs of 1 versus 2 technicians.

1. First, we note that $T = 2$ hours and $U = 20$ hours.
2. Then, $X = \dfrac{T}{T+U} = \dfrac{2}{2+20} = \dfrac{2}{22} = .091$ (close to .090).
3. For $M = 1$ server, $D = .350$ and $F = .960$.
4. For $M = 2$ servers, $D = .044$ and $F = .998$.
5. The average number of printers *working* is $J = NF(1 - X)$.
   For $M = 1$, this is $J = (5)(.960)(1 - .091) = 4.36$.
   For $M = 2$, it is $J = (5)(.998)(1 - .091) = 4.54$.
6. The cost analysis follows:

NUMBER OF TECHNICIANS	AVERAGE NUMBER PRINTERS DOWN $(N - J)$	AVERAGE COST/HR FOR DOWNTIME $(N - J)$(\$120/HR)	COST/HR. FOR TECHNICIANS (AT \$25/HR)	TOTAL COST/HR
1	.64	\$76.80	\$25.00	\$101.80
2	.46	\$55.20	\$50.00	\$105.20

This analysis suggests that having only one technician on duty will save a few dollars per hour (\$105.20 − \$101.80 = \$3.40).

[5]Limited, or finite, queuing tables are available to handle arrival populations of up to 250. Although there is no definite number that we can use as a dividing point between limited and unlimited populations, the general rule of thumb is this: If the number in the queue is a significant proportion of the arrival population, use a limited population queuing model. For a complete set of N-values, see L. G. Peck and R. N. Hazelwood, *Finite Queuing Tables* (New York: John Wiley, 1958).

TABLE D.8 ■ Finite Queuing Tables for a Population of $N = 5$

X	M	D	F	X	M	D	F	X	M	D	F	X	M	D	F	X	M	D	F
.012	1	.048	.999		1	.404	.945		1	.689	.801	.330	4	.012	.999		3	.359	.927
.019	1	.076	.998	.110	2	.065	.996	.210	3	.032	.998		3	.112	.986	.520	2	.779	.728
.025	1	.100	.997		1	.421	.939		2	.211	.973		2	.442	.904		1	.988	.384
.030	1	.120	.996	.115	2	.071	.995		1	.713	.783		1	.902	.583	.540	4	.085	.989
.034	1	.135	.995		1	.439	.933	.220	3	.036	.997	.340	4	.013	.999		3	.392	.917
.036	1	.143	.994	.120	2	.076	.995		2	.229	.969		3	.121	.985		2	.806	.708
.040	1	.159	.993		1	.456	.927		1	.735	.765		2	.462	.896		1	.991	.370
.042	1	.167	.992	.125	2	.082	.994	.230	3	.041	.997		1	.911	.569	.560	4	.098	.986
.044	1	.175	.991		1	.473	.920		2	.247	.965	.360	4	.017	.998		3	.426	.906
.046	1	.183	.990	.130	2	.089	.933		1	.756	.747		3	.141	.981		2	.831	.689
.050	1	.198	.989		1	.489	.914	.240	3	.046	.996		2	.501	.880		1	.993	.357
.052	1	.206	.988	.135	2	.095	.993		2	.265	.960		1	.927	.542	.580	4	.113	.984
.054	1	.214	.987		1	.505	.907		1	.775	.730	.380	4	.021	.998		3	.461	.895
.056	2	.018	.999	.140	2	.102	.992	.250	3	.052	.995		3	.163	.976		2	.854	.670
	1	.222	.985		1	.521	.900		2	.284	.955		2	.540	.863		1	.994	.345
.058	2	.019	.999	.145	3	.011	.999		1	.794	.712		1	.941	.516	.600	4	.130	.981
	1	.229	.984		2	.109	.991	.260	3	.058	.994	.400	4	.026	.977		3	.497	.883
.060	2	.020	.999		1	.537	.892		2	.303	.950		3	.186	.972		2	.875	.652
	1	.237	.983	.150	3	.012	.999		1	.811	.695		2	.579	.845		1	.996	.333
.062	2	.022	.999		2	.115	.990	.270	3	.064	.994		1	.952	.493	.650	4	.179	.972
	1	.245	.982		1	.553	.885		2	.323	.944	.420	4	.031	.997		3	.588	.850
.064	2	.023	.999	.155	3	.013	.999		1	.827	.677		3	.211	.966		2	.918	.608
	1	.253	.981		2	.123	.989	.280	3	.071	.993		2	.616	.826		1	.998	.308
.066	2	.024	.999		1	.568	.877		2	.342	.938		1	.961	.471	.700	4	.240	.960
	1	.260	.979	.160	3	.015	.999		1	.842	.661	.440	4	.037	.996		3	.678	.815
.068	2	.026	.999		2	.130	.988	.290	4	.007	.999		3	.238	.960		2	.950	.568
	1	.268	.978		1	.582	.869		3	.079	.992		2	.652	.807		1	.999	.286
.070	2	.027	.999	.165	3	.016	.999		2	.362	.932		1	.969	.451	.750	4	.316	.944
	1	.275	.977		2	.137	.987		1	.856	.644	.460	4	.045	.995		3	.763	.777
.075	2	.031	.999		1	.597	.861	.300	4	.008	.999		3	.266	.953		2	.972	.532
	1	.294	.973	.170	3	.017	.999		3	.086	.990		2	.686	.787	.800	4	.410	.924
.080	2	.035	.998		2	.145	.985		2	.382	.926		1	.975	.432		3	.841	.739
	1	.313	.969		1	.611	.853		1	.869	.628	.480	4	.053	.994		2	.987	.500
.085	2	.040	.998	.180	3	.021	.999	.310	4	.009	.999		3	.296	.945	.850	4	.522	.900
	1	.332	.965		2	.161	.983		3	.094	.989		2	.719	.767		3	.907	.702
.090	2	.044	.998		1	.638	.836		2	.402	.919		1	.980	.415		2	.995	.470
	1	.350	.960	.190	3	.024	.998		1	.881	.613	.500	4	.063	.992	.900	4	.656	.871
.095	2	.049	.997		2	.117	.980	.320	4	.010	.999		3	.327	.936		3	.957	.666
	1	.368	.955		1	.665	.819		3	.103	.988		2	.750	.748		2	.998	.444
.100	2	.054	.997	.200	3	.028	.998		2	.422	.912		1	.985	.399	.950	4	.815	.838
.100	1	.386	.950	.200	2	.194	.976		1	.892	.597	.520	4	.073	.991		3	.989	.631
.105	2	.059	.997																

OTHER QUEUING APPROACHES

Many practical waiting-line problems that occur in service systems have characteristics like those of the four mathematical models already described. Often, however, *variations* of these specific cases are present in an analysis. Service times in an automobile repair shop, for example, tend to follow the normal probability distribution instead of the exponential. A college registration system in which seniors have first choice of courses and hours over other students is an example of a first-come, first-served model with a preemptive priority queue discipline. A physical examination for military recruits is an example of a multiphase system, one that differs from the single-phase models discussed earlier in this module. A recruit first lines up to have blood drawn at one station, then waits for an eye exam at the next station, talks to a psychiatrist at the third, and is examined by a doctor for medical problems at the fourth. At each phase, the recruit must enter another queue and wait his or her turn. Many models, some very complex, have been developed to deal with situations such as these. One of these is described in the *OM in Action* box "Shortening Arraignment Times in New York's Police Department."

OM IN ACTION

Shortening Arraignment Times in New York's Police Department

At one time, people arrested in New York City averaged a 40-hour wait (some more than 70 hours) prior to arraignment. They were kept in crowded, noisy, stressful, unhealthy, and often dangerous holding facilities and, in effect, denied speedy court appearances. The New York Supreme Court has since ruled that the city must attempt to arraign within 24 hours or release the prisoner.

The arrest-to-arraignment (ATA) process, which has the general characteristics of a large queuing system, involves these steps: arrest of suspected criminal, transport to a police precinct, search/fingerprinting, paperwork for arrest, transport to a central booking facility, additional paperwork, processing of fingerprints, a bail interview, transport to either the courthouse or an outlying precinct, checks for a criminal record, and finally, drawing up of a complaint document by an assistant district attorney.

To solve the complex problem of improving this system, the city hired Queues Enforth Development, Inc., a Massachusetts consulting firm. The firm's Monte Carlo simulation of the ATA process included single- and multiple-server queuing models. The modeling approach successfully reduced the average ATA time to 24 hours and resulted in an annual cost savings of $9.5 million for the city and state.

Source: R. C. Larson, M. F. Cahn, and M. C. Shell, "Improving the New York Arrest-to-Arraignment System," *Interfaces* 23, no. 1 (January–February 1993):76–96.

SUMMARY

Queues are an important part of the world of operations management. In this module, we describe several common queuing systems and present mathematical models for analyzing them.

The most widely used queuing models include Model A, the basic single-channel, single-phase system with Poisson arrivals and exponential service times; Model B, the multichannel equivalent of Model A; Model C, a constant-service-rate model; and Model D, a limited-population system. All four models allow for Poisson arrivals, first-in, first-out service, and a single-service phase. Typical operating characteristics we examine include average time spent waiting in the queue and system, average number of customers in the queue and system, idle time, and utilization rate.

A variety of queuing models exists for which all the assumptions of the traditional models need not be met. In these cases, we use more complex mathematical models or turn to a technique called *simulation*. The application of simulation to problems of queuing systems is addressed in Quantitative Module F.

KEY TERMS

Queuing theory *(p. 744)*
Waiting line (queue) *(p. 744)*
Unlimited, or infinite, population *(p. 745)*
Limited, or finite, population *(p. 745)*
Poisson distribution *(p. 746)*
First-in, first-out (FIFO) rule *(p. 747)*
Single-channel queuing system *(p. 747)*
Multiple-channel queuing system *(p. 747)*
Single-phase system *(p. 747)*
Multiphase system *(p. 747)*
Negative exponential probability distribution *(p. 747)*

USING SOFTWARE TO SOLVE QUEUING PROBLEMS

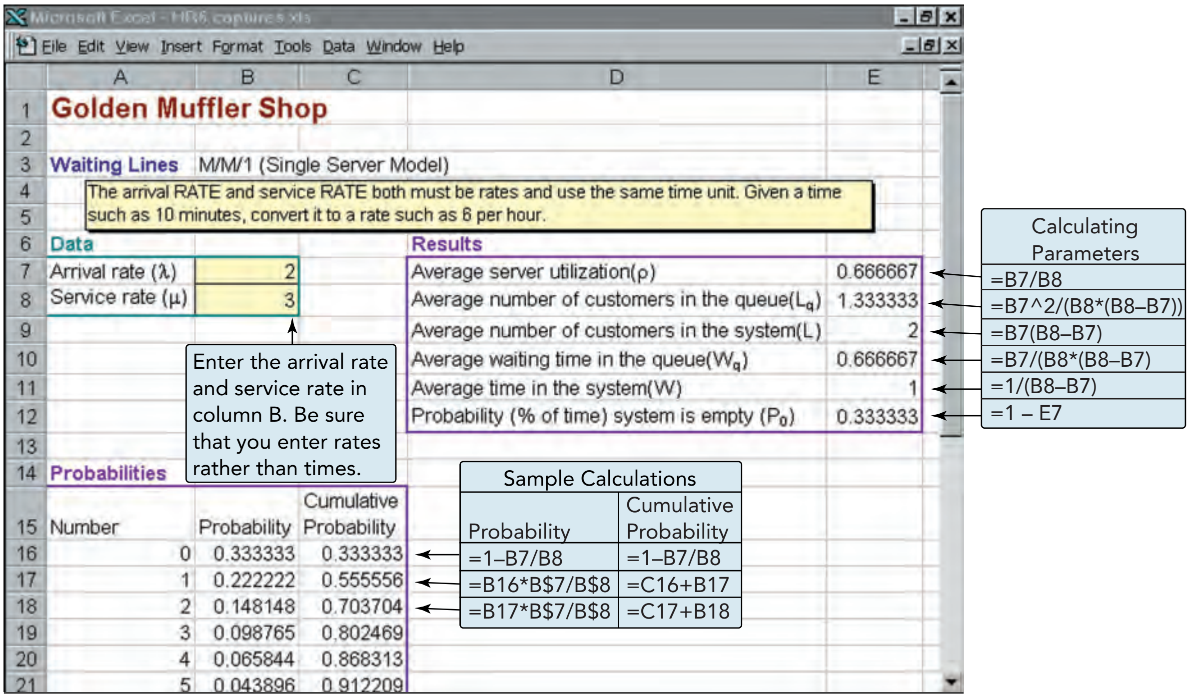
Both Excel OM and POM for Windows may be used to analyze all but the last two homework problems in this module.

Using Excel OM

Excel OM's Waiting-Line program handles all four of the models developed in this module. Program D.1 illustrates our first model, the M/M/1 system, using the data from Example D1.

Using POM For Windows

There are several POM for Windows queuing models from which to select in that program's Waiting-Line module. The program can include an economic analysis of cost data, and, as an option, you may display probabilities of various numbers of people/items in the system. See Appendix IV for further details.



PROGRAM D.1 ■ Using Excel OM for Queuing

Example D1's (Golden Muffler Shop) data are illustrated in the M/M/1 model.

SOLVED PROBLEMS

Solved Problem D.1

Sid Das Brick Distributors currently employs 1 worker whose job is to load bricks on outgoing company trucks. An average of 24 trucks per day, or 3 per hour, arrive at the loading platform, according to a Poisson distribution. The worker loads them at a rate of 4 trucks per hour, following approximately the exponential distribution in his service times.

Das believes that adding a second brick loader will substantially improve the firm's productivity. He estimates that a two-person crew at the loading gate will double the loading rate from 4 trucks per hour to 8 trucks per hour. Analyze the effect on the queue of such a change and compare the results to those achieved with one worker. What is the probability that there will be more than 3 trucks either being loaded or waiting?

SOLUTION

	NUMBER OF BRICK LOADERS	
	1	2
Truck arrival rate (λ)	3/hr.	3/hr.
Loading rate (μ)	4/hr.	8/hr.
Average number in system (L_s)	3 trucks	.6 truck
Average time in system (W_s)	1 hr.	.2 hr.
Average number in queue (L_q)	2.25 trucks	.225 truck
Average time in queue (W_q)	.75 hr.	.075 hr.
Utilization rate (ρ)	.75	.375
Probability system empty (P_0)	.25	.625

Probability of More Than k Trucks in System

	PROBABILITY $n > k$	
k	ONE LOADER	TWO LOADERS
0	.75	.375
1	.56	.141
2	.42	.053
3	.32	.020

These results indicate that when only one loader is employed, the average truck must wait three quarters of an hour before it is loaded. Furthermore, there is an average of 2.25 trucks waiting in line to be loaded. This situation may be unacceptable to management. Note also the decline in queue size after the addition of a second loader.

Solved Problem D.2

Truck drivers working for Sid Das (see Solved Problem D.1) earn an average of $10 per hour. Brick loaders receive about $6 per hour. Truck drivers waiting *in the queue or at the loading platform* are drawing a salary but are productively idle and unable to generate revenue during that time. What would be the *hourly* cost savings to the firm if it employed 2 loaders instead of 1?

Referring to the data in Solved Problem D.1, we note that the average number of trucks *in the system* is 3 when there is only 1 loader and .6 when there are 2 loaders.

SOLUTION

	NUMBER OF LOADERS	
	1	**2**
Truck driver idle time costs [(Average number of trucks) × (Hourly rate)] = (3)($10) =	$30	$ 6 = (.6)($10)
Loading costs	6	12 = (2)($6)
Total expected cost per hour	$36	$18

The firm will save $18 per hour by adding a second loader.

Solved Problem D.3

Sid Das is considering building a second platform or gate to speed the process of loading trucks. This system, he thinks, will be even more efficient than simply hiring another loader to help out on the first platform (as in Solved Problem D.1).

Assume that workers at each platform will be able to load 4 trucks per hour each and that trucks will continue to arrive at the rate of 3 per hour. Then apply the appropriate equations to find the waiting line's new operating conditions. Is this new approach indeed speedier than the other two that Das has considered?

SOLUTION

$$P_0 = \frac{1}{\left[\displaystyle\sum_{n=0}^{1} \frac{1}{n!}\left(\frac{3}{4}\right)^n\right] + \frac{1}{2!}\left(\frac{3}{4}\right)^2 \frac{2(4)}{2(4)-3}}$$

$$= \frac{1}{1 + \frac{3}{4} + \frac{1}{2}\left(\frac{3}{4}\right)^2\left(\frac{8}{8-3}\right)} = .454$$

$$L_s = \frac{3(4)(3/4)^2}{(1)!\,(8-3)^2}(.4545) + \frac{3}{4} = .873$$

$$W_s = \frac{.873}{3} = .291 \text{ hr.}$$

$$L_q = .873 - 3/4 = .123$$

$$W_q = \frac{.123}{3} = 0.41 \text{ hr.}$$

Looking back at Solved Problem D.1, we see that although length of the *queue* and average time in the queue are lowest when a second platform is open, the average number of trucks in the *system* and average time spent waiting in the system are smallest when two workers are employed at a *single* platform. Thus, we would probably recommend not building a second platform.

Solved Problem D.4

St. Elsewhere Hospital's Cardiac Care Unit (CCU) has 5 beds, which are virtually always occupied by patients who have just undergone major heart surgery. Two registered nurses are on duty in the CCU in each of the three 8-hour shifts. About every 2 hours (following a Poisson distribution), one of the patients requires a nurse's attention. The nurse will then spend an average of 30 minutes (exponentially distributed) assisting the patient and updating medical records regarding the problem and care provided.

Because immediate service is critical to the 5 patients, two important questions are: What is the average number of patients being attended by the nurses? What is the average time that a patient spends waiting for one of the nurses to arrive?

SOLUTION

N = 5 patients

M = 2 nurses

T = 30 minutes

U = 120 minutes

$$X = \frac{T}{T+U} = \frac{30}{30+120} = .20$$

From Table D.8 (p. 758), with $X = .20$ and $M = 2$, we see that

$$F = .976$$

$$H = \text{average number being attended to} = FNX$$
$$= (.976)(5)(.20) = .98 \approx 1 \text{ patient at any given time}$$

$$W = \text{average waiting time for a nurse} = \frac{T(1-F)}{XF}$$
$$= \frac{30(1-.976)}{(.20)(.976)} = 3.69 \text{ minutes}$$

INTERNET AND STUDENT CD-ROM EXERCISES

Visit our Companion Web site or use your student CD-ROM to help with material in this module.

On Our Companion Web site, www.prenhall.com/heizer

- Self-Study Quizzes
- Practice Problems
- Internet Homework Problems
- Internet Case

On Your Student CD-ROM

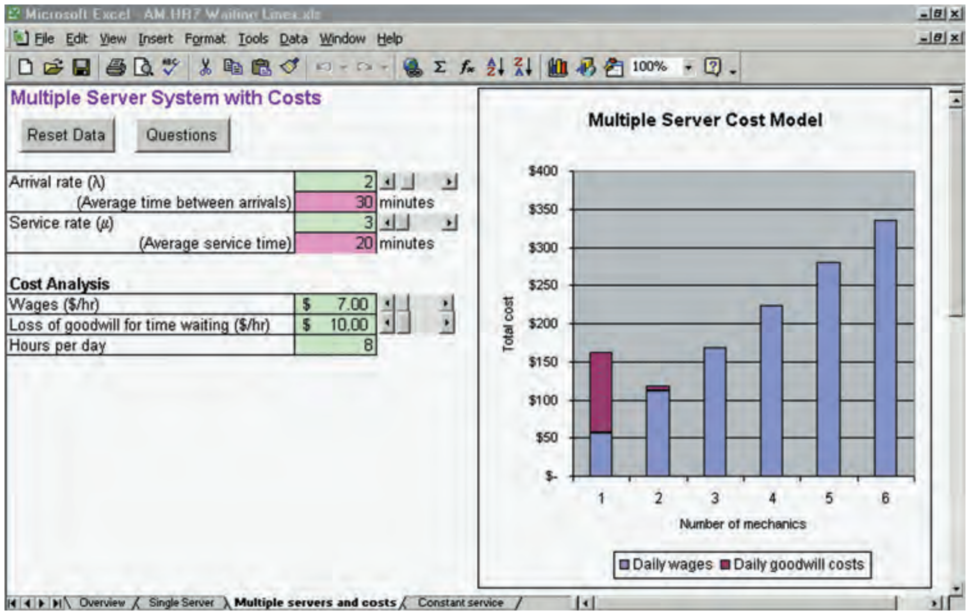
- PowerPoint Lecture
- Practice Problems
- Active Model Exercises
- Excel OM
- Excel OM Example Data Files
- POM for Windows

DISCUSSION QUESTIONS

1. Name the three parts of a typical queuing system.
2. When designing a waiting line system, what "qualitative" concerns need to be considered?
3. Name the three factors that govern the structure of "arrivals" in a queuing system.
4. State the seven common measures of queuing system performance.
5. State the assumptions of the "basic" single-channel queuing model (Model A or M/M/1).
6. Is it good or bad to operate a supermarket bakery system on a strict first-come, first-served basis? Why?
7. Describe what is meant by the waiting-line terms *balk* and *renege*. Provide an example of each.
8. Which is larger, W_s or W_q? Explain.
9. Briefly describe three situations in which the first-in, first-out (FIFO) discipline rule is not applicable in queuing analysis.
10. Describe the behavior of a waiting line where $\lambda > \mu$. Use both analysis and intuition.
11. Discuss the likely outcome of a waiting line system where $\mu > \lambda$, but only by a tiny amount. (For example, $\mu = 4.1$, $\lambda = 4$).
12. Provide examples of four situations in which there is a limited, or finite, waiting line.
13. What are the components of the following queuing systems? Draw and explain the configuration of each.
    (a) Barbershop.
    (b) Car wash.
    (c) Laundromat.
    (d) Small grocery store.
14. Do doctors' offices generally have random arrival rates for patients? Are service times random? Under what circumstances might service times be constant?
15. What happens if two single-channel systems have the same mean arrival and service rates, but the service time is constant in one and exponential in the other?
16. What dollar value do you place on yourself per hour that you spend waiting in lines? What value do your classmates place on themselves? Why do the values differ?

ACTIVE MODEL EXERCISE

In this active model example, we can examine the relationship between arrival and service rates, and costs and the number of servers. The first two inputs to the model are arrival and service rates in customers per hour. The average time between arrivals and service time is also displayed, but do not change these two numbers, which are shaded in pink.



ACTIVE MODEL D.2

An Analysis of the Golden Muffler Shop (Examples D1–D3) with Cost as a Variable

Questions

1. What number of mechanics yields the lowest total daily cost? What is the minimum total daily cost?
2. Use the scrollbar on the arrival rate. What would the arrival rate need to be to require a third mechanic?
3. Use the scrollbar on the goodwill cost and determine the range of goodwill costs for which you would have exactly 1 mechanic. Two mechanics?
4. How high would the wage rate need to be to make 1 mechanic the least costly option?
5. If a second mechanic is added, is it less costly to have the 2 mechanics working separately or to have the 2 mechanics work as a single team with a service rate that is twice as fast?

PROBLEMS*

• **D.1** Customers arrive at Paul Harrold's Styling Shop at a rate of 3 per hour, distributed in a Poisson fashion. Paul can perform haircuts at a rate of 5 per hour, distributed exponentially.

a) Find the average number of customers waiting for haircuts.
b) Find the average number of customers in the shop.
c) Find the average time a customer waits until it is his or her turn.
d) Find the average time a customer spends in the shop.
e) Find the percentage of time that Paul is busy.

• **D.2** There is only one copying machine in the student lounge of the business school. Students arrive at the rate of $\lambda = 40$ per hour (according to a Poisson distribution). Copying takes an average of 40 seconds, or $\mu = 90$ per hour (according to an exponential distribution). Compute the following:

a) The percentage of time that the machine is used.
b) The average length of the queue.
c) The average number of students in the system.
d) The average time spent waiting in the queue.
e) The average time in the system.

*Note: **P** means the problem may be solved with POM for Windows; [X] means the problem may be solved with Excel OM; and **P**[X] means the problem may be solved with POM for Windows and/or Excel OM.

• **D.3** Glen Schmidt owns and manages a chili-dog and soft-drink stand near the Georgeville campus. While Glen can service 30 customers per hour on the average (μ), he gets only 20 customers per hour (λ). Because Glen could wait on 50% more customers than actually visit his stand, it doesn't make sense to him that he should have any waiting lines.

Glen hires you to examine the situation and to determine some characteristics of his queue. After looking into the problem, you find it follows the six conditions for a single-channel waiting line (as seen in Model A). What are your findings?

• **D.4** Sam Certo, a Longwood vet, is running a rabies vaccination clinic for dogs at the local grade school. Sam can "shoot" a dog every 3 minutes. It is estimated that the dogs will arrive independently and randomly throughout the day at a rate of one dog every 6 minutes according to a Poisson distribution. Also assume that Sam's shooting times are exponentially distributed. Compute the following:

a) The probability that Sam is idle.
b) The proportion of the time that Sam is busy.
c) The average number of dogs being vaccinated and waiting to be vaccinated.
d) The average number of dogs waiting to be vaccinated.
e) The average time a dog waits before getting vaccinated.
f) The average amount of time a dog spends waiting in line and being vaccinated.

: **D.5** The pharmacy at Arnold Palmer Hospital receives 12 requests for prescriptions each hour, Poisson distributed. It takes the staff a mean time of 4 minutes to fill each, following a negative exponential distribution.

a) What is the number of units in the system (waiting time plus service time).
b) How long will the typical (average) prescription be in the system (waiting plus service time)?
c) What is the average number of prescriptions in the queue?

• **D.6** Calls arrive at James Hamann's hotel switchboard at a rate of 2 per minute. The average time to handle each is 20 seconds. There is only one switchboard operator at the current time. The Poisson and exponential distributions appear to be relevant in this situation.

a) What is the probability that the operator is busy?
b) What is the average time that a caller must wait before reaching the operator?
c) What is the average number of calls waiting to be answered?

: **D.7** Automobiles arrive at the drive-through window at the downtown Urbana, Illinois, post office at the rate of 4 every 10 minutes. The average service time is 2 minutes. The Poisson distribution is appropriate for the arrival rate and service times are exponentially distributed.

a) What is the average time a car is in the system?
b) What is the average number of cars in the system?
c) What is the average number of cars waiting to receive service?
d) What is the average time a car is in the queue?
e) What is probability that there are no cars at the window?
f) What percentage of the time is the postal clerk busy?
g) What is the probability that there are exactly 2 cars in the system?

• **D.8** The Tara Yazinski Electronics Corporation retains a service crew to repair machine breakdowns that occur on an average of $\lambda = 3$ per 8-hour workday (approximately Poisson in nature). The crew can service an average of $\mu = 8$ machines per workday, with a repair time distribution that resembles the exponential distribution.

a) What is the utilization rate of this service system?
b) What is the average downtime for a broken machine?
c) How many machines are waiting to be serviced at any given time?
d) What is the probability that more than one machine is in the system? The probability that more than two are broken and waiting to be repaired or being serviced? More than three? More than four?

: **D.9** Zimmerman's Bank is the only bank in the small town of St. Thomas. On a typical Friday, an average of 10 customers per hour arrive at the bank to transact business. There is one teller at the bank, and the average time required to transact business is 4 minutes. It is assumed that service times may be described by the exponential distribution. A single line would be used, and the customer at the front of the line would go to the first available bank teller. If a single teller is used, find:

a) The average time in the line.
b) The average number in the line.
c) The average time in the system.
d) The average number in the system.
e) The probability that the bank is empty.
f) Zimmerman is considering adding a second teller (who would work at the same rate as the first) to reduce the waiting time for customers. She assumes that this will cut the waiting time in half. If a second teller is added, find the new answers to parts (a) to (e).

D.10 Valerie Fondl manages a Columbus, Ohio, movie theater complex called Cinema I, II, III, and IV. Each of the four auditoriums plays a different film; the schedule staggers starting times to avoid the large crowds that would occur if all four movies started at the same time. The theater has a single ticket booth and a cashier who can maintain an average service rate of 280 patrons per hour. Service times are assumed to follow an exponential distribution. Arrivals on a normally active day are Poisson distributed and average 210 per hour.

To determine the efficiency of the current ticket operation, Valeri wishes to examine several queue-operating characteristics.

a) Find the average number of moviegoers waiting in line to purchase a ticket.
b) What percentage of the time is the cashier busy?
c) What is the average time that a customer spends in the system?
d) What is the average time spent waiting in line to get to the ticket window?
e) What is the probability that there are more than two people in the system? More than three people? More than four?

D.11 Bill Youngdahl has been collecting data at the TU student grill. He has found that, between 5:00 P.M. and 7:00 P.M., students arrive at the grill at a rate of 25 per hour (Poisson distributed) and service time takes an average of 2 minutes (exponential distribution). There is only 1 server, who can work on only 1 order at a time.

a) What is the average number of students in line?
b) What is the average time a student is in the grill area?
c) Suppose that a second server can be added to team up with the first (and, in effect, act as one faster server). This would reduce the average service time to 90 seconds. How would this affect the average time a student is in the grill area?
d) Suppose a second server is added and the 2 servers act independently, with *each* taking an average of 2 minutes. What would be the average time a student is in the system?

D.12 The wheat harvesting season in the American Midwest is short, and farmers deliver their truckloads of wheat to a giant central storage bin within a 2-week span. Because of this, wheat-filled trucks waiting to unload and return to the fields have been known to back up for a block at the receiving bin. The central bin is owned cooperatively, and it is to every farmer's benefit to make the unloading/storage process as efficient as possible. The cost of grain deterioration caused by unloading delays and the cost of truck rental and idle driver time are significant concerns to the cooperative members. Although farmers have difficulty quantifying crop damage, it is easy to assign a waiting and unloading cost for truck and driver of $18 per hour. During the 2-week harvest season, the storage bin is open and operated 16 hours per day, 7 days per week, and can unload 35 trucks per hour according to an exponential distribution. Full trucks arrive all day long (during the hours the bin is open) at a rate of about 30 per hour, following a Poisson pattern.

To help the cooperative get a handle on the problem of lost time while trucks are waiting in line or unloading at the bin, find the following:

a) The average number of trucks in the unloading system.
b) The average time per truck in the system.
c) The utilization rate for the bin area.
d) The probability that there are more than three trucks in the system at any given time.
e) The total daily cost to the farmers of having their trucks tied up in the unloading process.
f) As mentioned, the cooperative uses the storage bin heavily only 2 weeks per year. Farmers estimate that enlarging the bin would cut unloading costs by 50% next year. It will cost $9,000 to do so during the off-season. Would it be worth the expense to enlarge the storage area?

D.13 Radovilsky's Department Store in Haywood, California, maintains a successful catalog sales department in which a clerk takes orders by telephone. If the clerk is occupied on one line, incoming phone calls to the catalog department are answered automatically by a recording machine and asked to wait. As soon as the clerk is free, the party who has waited the longest is transferred and serviced first. Calls come in at a rate of about 12 per hour. The clerk can take an order in an average of 4 minutes. Calls tend to follow a Poisson distribution, and service times tend to be exponential.

The cost of the clerk is $10 per hour, but because of lost goodwill and sales, Radovilsky's loses about $25 per hour of customer time spent waiting for the clerk to take an order.

a) What is the average time that catalog customers must wait before their calls are transferred to the order clerk?
b) What is the average number of callers waiting to place an order?
c) Radovilsky's is considering adding a second clerk to take calls. The store's cost would be the same $10 per hour. Should it hire another clerk? Explain your decision.

D.14 Customers arrive at an automated coffee-vending machine at a rate of four per minute, following a Poisson distribution. The coffee machine dispenses cups of coffee at a constant time of 10 seconds.

a) What is the average number of people waiting in line?
b) What is the average number in the system?
c) How long does the average person wait in line before receiving service?

D.15 The typical subway station in Washington, DC, has six turnstiles, each of which can be controlled by the station manager to be used for either entrance or exit control—but never for both. The manager must decide at different times of the day how many turnstiles to use for entering passengers and how many to use for exiting passengers.

At the George Washington University (GWU) Station, passengers enter the station at a rate of about 84 per minute between the hours of 7 A.M. and 9 A.M. Passengers exiting trains at the stop reach the exit turnstile area at a rate of about 48 per minute during the same morning rush hours. Each turnstile can allow an average of 30 passengers per minute to enter or exit. Arrival and service times have been thought to follow Poisson and exponential distributions, respectively. Assume riders form a common queue at both entry and exit turnstile areas and proceed to the first empty turnstile.

The GWU station manager, Ernie Forman, does not want the average passenger at his station to have to wait in a turnstile line for more than 6 seconds, nor does he want more than 8 people in any queue at any average time.

a) How many turnstiles should be opened in each direction every morning?
b) Discuss the assumptions underlying the solution of this problem using queuing theory.

D.16 Yvette Freeman's Car Wash takes a constant time of 4.5 minutes in its automated car wash cycle. Autos arrive following a Poisson distribution at the rate of 10 per hour. Yvette wants to know:

a) The average waiting time in line.
b) The average length of the line.

D.17 Eric Krassow's cabinet-making shop, in Memphis, has five tools that automate the drilling of holes for the installation of hinges. These machines need setting up for each order of cabinets. The orders appear to follow the Poisson distribution, averaging 3 per 8-hour day. There is a single technician for setting these machines. His service times are exponential, averaging 2 hours each.

a) What is the service factor for this system?
b) What is the average number of these machines in service?
c) What impact on machines in service would there be if a second technician were available?

D.18 Two technicians, working separately, monitor a group of 5 computers that run an automated manufacturing facility. It takes an average of 15 minutes (exponentially distributed) to adjust a computer that develops a problem. Computers run for an average of 85 minutes (Poisson distributed) without requiring adjustments. Determine the following:

a) The average number of computers waiting for adjustment.
b) The average number being adjusted.
c) The average number of computers not in working order.

D.19 One mechanic services 5 drilling machines for a steel plate manufacturer. Machines break down on an average of once every 6 working days, and breakdowns tend to follow a Poisson distribution. The mechanic can handle an average of one repair job per day. Repairs follow an exponential distribution.

a) On the average, how many machines are waiting for service?
b) On the average, how many drills are in running order?
c) How much would waiting time be reduced if a second mechanic were hired?

D.20 The administrator at a large hospital emergency room faces the problem of providing treatment for patients who arrive at different rates during the day. There are 4 doctors available to treat patients when needed. If not needed, they can be assigned other responsibilities (such as doing lab tests, reports, X-ray diagnoses) or else rescheduled to work at other hours.

It is important to provide quick and responsive treatment, and the administrator feels that, on the average, patients should not have to sit in the waiting area for more than 5 minutes before being seen by a doctor. Patients are treated on a first-come, first-served basis and see the first available doctor after waiting in the queue. The arrival pattern for a typical day is as follows:

TIME	ARRIVAL RATE
9 A.M.–3 P.M.	6 patients/hour
3 P.M.–8 P.M.	4 patients/hour
8 P.M.–midnight	12 patients/hour

Arrivals follow a Poisson distribution, and treatment times, 12 minutes on the average, follow the exponential pattern.

How many doctors should be on duty during each period to maintain the level of patient care expected?

D.21 The Chattanooga Furniture store gets an average of 50 customers per shift. Marilyn Helms, the manager, wants to calculate whether she should hire 1, 2, 3, or 4 salespeople. She has determined that average waiting times will be 7 minutes with one salesperson, 4 minutes with two salespeople, 3 minutes with three salespeople, and 2 minutes with four salespeople. She has estimated the cost per minute that customers wait at \$1. The cost per salesperson per shift (including fringe benefits) is \$70.

How many salespeople should be hired?

D.22 Gather real arrival and service data from somewhere on campus or some other locale (perhaps a bank, barbershop, car wash, or fast-food restaurant). Then, address the following questions:

a) Determine the distributions of the arrivals and the service times (plot both).
b) Did the arrivals or the services follow the distributions discussed in the text? (Plot these distributions on the same graph as your raw data.)
c) What queuing model did your "real" queue follow?
d) Determine the average length of the queue.
e) Determine the average number of customers in the system.

INTERNET HOMEWORK PROBLEMS

See our Companion Web site at **www.prenhall.com/heizer** for these additional homework problems: D.23 through D.31.

CASE STUDY

New England Foundry

For more than 75 years, New England Foundry, Inc. (NEFI), has manufactured wood stoves for home use. In recent years, with increasing energy prices, president George Mathison has seen sales triple. This dramatic increase has made it difficult for George to maintain quality in all his wood stoves and related products.

Unlike other companies manufacturing wood stoves, NEFI is in the business of making *only* stoves and stove-related products. Its major products are the Warmglo I, the Warmglo II, the Warmglo III, and the Warmglo IV. The Warmglo I is the smallest wood stove, with a heat output of 30,000 BTUs, and the Warmglo IV is the largest, with a heat output of 60,000 BTUs.

The Warmglo III outsold all other models by a wide margin. Its heat output and available accessories were ideal for the typical home. The Warmglo III also had a number of other outstanding features that made it one of the most attractive and heat-efficient stoves on the market. These features, along with the accessories, resulted in expanding sales and prompted George to build a new factory to manufacture the Warmglo III model. An overview diagram of the factory is shown in Figure D.6.

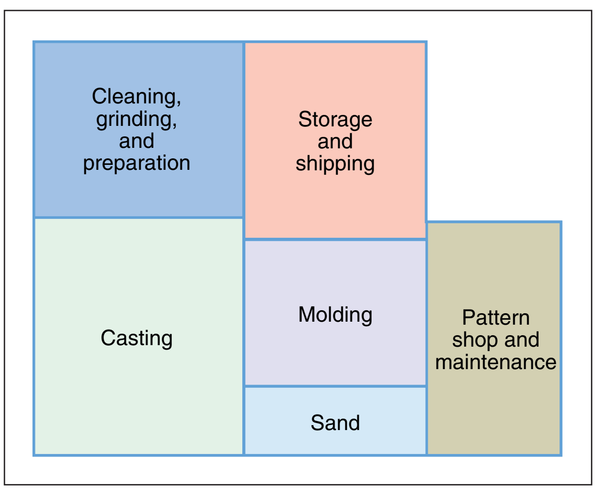


FIGURE D.6 ■ Overview of Factory

The new foundry used the latest equipment, including a new Disamatic that helped in manufacturing stove parts. Regardless of new equipment or procedures, casting operations have remained basically unchanged for hundreds of years. To begin with, a wooden pattern is made for every cast-iron piece in the stove. The wooden pattern is an exact duplicate of the cast-iron piece that is to be manufactured. All NEFI patterns are made by Precision Patterns, Inc. and are stored in the pattern shop and maintenance room. Next, a specially formulated sand is molded around the wooden pattern. There can be two or more sand molds for each pattern. The sand is mixed and the molds are made in the molding room. When the wooden pattern is removed, the resulting sand molds form a negative image of the desired casting. Next, molds are transported to the casting room, where molten iron is poured into them and allowed to cool. When the iron has solidified, molds are moved into the cleaning, grinding, and preparation room, where they are dumped into large vibrators that shake most of the sand from the casting. The rough castings are then subjected to both sandblasting to remove the rest of the sand and grinding to finish some of their surfaces. Castings are then painted with a special heat-resistant paint, assembled into workable stoves, and inspected for manufacturing defects that may have gone undetected. Finally, finished stoves are moved to storage and shipping, where they are packaged and transported to the appropriate locations.

At present, the pattern shop and the maintenance department are located in the same room. One large counter is used by both maintenance personnel, who store tools and parts (which are mainly used by the casting department); and sand molders, who need various patterns for the molding operation. Pete Nawler and Bob Dillman, who work behind the counter, can service a total of 10 people per hour (about 5 per hour each). On the average, 4 people from casting and 3 from molding arrive at the counter each hour. People from molding and casting departments arrive randomly, and to be served, they form a single line.

Pete and Bob have always had a policy of first come, first served. Because of the location of the pattern shop and maintenance department, it takes an average of 3 minutes for an individual from the casting department to walk to the pattern and maintenance room, and it takes about 1 minute for an individual to walk from the molding department to the pattern and maintenance room.

After observing the operation of the pattern shop and maintenance room for several weeks, George decided to make some changes to the factory layout. An overview of these changes appears in Figure D.7.

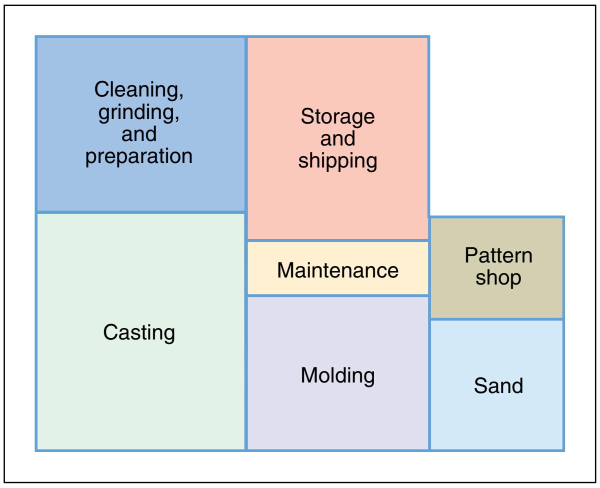


FIGURE D.7 ■ Overview of Factory after Changes

Separating the maintenance shop from the pattern shop had a number of advantages. It would take people from the casting department only 1 minute instead of 3 to get to the new maintenance room. The time from molding to the pattern shop would be unchanged. Using motion and time studies, George was also able to determine that improving the layout of the maintenance room would allow Bob to serve 6 people from the casting department per hour; improving the layout of the pattern department would allow Pete to serve 7 people from the molding shop per hour.

Discussion Questions

1. How much time would the new layout save?
2. If casting personnel were paid $9.50 per hour and molding personnel were paid $11.75 per hour, how much could be saved per hour with the new factory layout?
3. Should George have made the change in layout?

Source: From *Quantitative Analysis for Management*, 9th ed., by B. Render, R. Stair, and M. Hanna pp. 602–604. Copyright © 2006. Reprinted by permission of Prentice Hall, Inc., Upper Saddle River, NJ.

CASE STUDY

The Winter Park Hotel

Donna Shader, manager of the Winter Park Hotel, is considering how to restructure the front desk to reach an optimum level of staff efficiency and guest service. At present, the hotel has five clerks on duty, each with a separate waiting line, during peak check-in time of 3:00 P.M. to 5:00 P.M. Observation of arrivals during this period shows that an average of 90 guests arrive each hour (although there is no upward limit on the number that could arrive at any given time). It takes an average of 3 minutes for the front-desk clerk to register each guest.

Ms. Shader is considering three plans for improving guest service by reducing the length of time that guests spend waiting in line. The first proposal would designate one employee as a quick-service clerk for guests registering under corporate accounts, a market segment that fills about 30% of all occupied rooms. Because corporate guests are preregistered, their registration takes just 2 minutes. With these guests separated from the rest of the clientele, the average time for registering a typical guest would climb to 3.4 minutes. Under this plan, noncorporate guests would choose any of the remaining four lines.

The second plan is to implement a single-line system. All guests could form a single waiting line to be served by whichever of five clerks became available. This option would require sufficient lobby space for what could be a substantial queue.

The use of an automatic teller machine (ATM) for check-ins is the basis of the third proposal. This ATM would provide about the same service rate as would a clerk. Because initial use of this technology might be minimal, Shader estimates that 20% of customers, primarily frequent guests, would be willing to use the machines. (This might be a conservative estimate if guests perceive direct benefits from using the ATM, as bank customers do. Citibank reports that some 95% of its Manhattan customers use its ATMs.) Ms. Shader would set up a single queue for customers who prefer human check-in clerks. This line would be served by the five clerks, although Shader is hopeful that the ATM machine will allow a reduction to four.

Discussion Questions

1. Determine the average amount of time that a guest spends checking in. How would this change under each of the stated options?
2. Which option do you recommend?

Source: From *Quantitative Analysis for Management*, 9th ed., by B. Render, R. Stair, and M. Hanna. p. 604. Copyright © 2006. Reprinted by permission of Prentice Hall, Inc., Upper Saddle River, NJ.

ADDITIONAL CASE STUDY

See our Companion Web site at www.prenhall.com/heizer for this additional free case study:

- **Pantry Shopper**: The case requires the redesign of a checkout system for a supermarket.

BIBLIOGRAPHY

Barron, K. "Hurry Up and Wait," *Forbes* (October 16, 2000): 158–164.

Bennett, J. C., and D. J. Worthington. "An Example of Good but Partially Successful OR Engagement: Improving Outpatient Clinic Operations," *Interfaces* 28, no. 5 (September–October 1998): 56–69.

Dasgupta, Ani, and Ghosh, Madhubani. "Including Performance in a Queue via Prices: The Case of a Riverine Port." *Management Science* 46, no. 11 (November 2000): 1466–1484.

Haksever, C., B. Render, and R. Russell. *Service Management and Operations*, 2nd ed. Upper Saddle River, NJ: Prentice Hall (2000).

Mandelbaum, A., and M. I. Reiman. "On Pooling Queuing Networks." *Management Science* 44, no. 7 (July 1998): 971–981.

Prabhu, N. U. *Foundations of Queuing Theory*. Dordecht, Netherlands: Kluwer Academic Publishers (1997).

Render, B., R. M. Stair, and R. Balakrishnan. *Managerial Decision Modeling with Spreadsheets*, 2nd ed. Upper Saddle River, NJ: Prentice Hall (2006).

Render, B., R. M. Stair, and M. Hanna. *Quantitative Analysis for Management*, 9th ed. Upper Saddle River, NJ: Prentice Hall (2006).

Samuelson, Douglas A. "Predictive Dialing for Outbound Telephone Call Centers." *Interfaces* 29, no. 5 (September 1999): 66–81.

Whitt, W. "Predicting Queuing Delays." *Management Science* 45, no. 6 (June 1999): 870–888.

INTERNET RESOURCES

Journal: Methodology and Computing in Applied Probability:
http://www.maths.uq.edu.au/~pkp/misc/mcap.html

Notes on Queuing Theory:
http://mscmga.ms.ic.ac.uk/jeb/or/queue.html