Chapter Outline

Analyzing Queues Using Analytical Models
- Arrival Distribution
- Service-Time Distribution
- Queue Discipline
- Queuing Behavior

Single-Server Queuing Model
Multiple-Server Queuing Model
The Economics of Waiting-Line Analysis
OM Spotlight: Airport Security Wait Times

The Psychology of Waiting
- Solved Problems
- Key Terms and Concepts
- Questions for Review and Discussion
- Problems and Activities
- Cases
  - Bourbon County Court
- Endnotes

Learning Objectives

- To understand the key elements and underlying mathematical concepts of analytical queuing models: arrival distribution, service-time distribution, queue discipline, and queuing behavior.
- To understand the single-server queuing model and be able to calculate and interpret the operating characteristics associated with the model.
- To be able to apply the operating characteristic formulas for a multiple-server queuing model to evaluate performance of practical queuing systems.
- To understand economic trade-offs associated with designing and managing queuing systems.
- To appreciate the importance of understanding the psychology of waiting in designing and managing queuing systems in practical business situations.
• An electrical utility company uses six customer service representatives (CSRs) at its call center to handle telephone calls and inquiries from its top 350 business customers. The next tier of 700 business customers is also handled by six CSRs. Based on the customer’s code, the call center routes business customers to different queues and CSRs. A manager at the utility explains: “We don’t ignore anyone, but our biggest customers certainly get more attention than the rest.”¹

• Amtrak, the U.S. passenger train service, pays freight railroads a fee to use their tracks. With freight train business growing, Amtrak trains were on time only 63 percent of the time in July, 2004. When freight trains back up, there is a spillover effect on passenger trains. Amtrak’s Sunset Limited passenger train between Orlando, Florida and Los Angeles hasn’t been on time once in the previous five months. Amtrak terminated one trip in El Paso, Texas when the Sunset Limited arrived 35 hours late—it forced Amtrak to bus and fly all passengers between El Paso and points West. Amtrak passenger traffic is up 6%, in part because more airline travelers are fed up with delays at the airports.²

The first episode highlights a growing practice of segmenting customers so that premium service is provided to a few high-value customers while many low-value customers get less attention and organizational resources. The electric utility’s call center assigns the same number of CSRs—six—to the top 350 customers and the next 700 customers based on value. Many organizations would gladly see customers that generate marginal profits leave. Value-based queuing is a method that allows organizations to prioritize customer calls based on their long-term value to the organization. Low-profitability customers are often encouraged to serve themselves on the company’s web site rather than tie up expensive telephone representatives. Such decisions are similar to the notion of segmenting high-value inventory using ABC analysis that we discussed in Chapter 12. Do you think that this decision is good or bad? Should all customers be treated the same and be considered as important as any other?

The second episode illustrates the interdependency between U.S. airline, passenger train, and freight train networks. Major delays in one transportation system can quickly cascade to other types of transportation with huge cost implications. For all three types of transportation, the management of routes, schedules, queues, capacity, and the use of priority rules are critical to the efficiency of these interdependent transportation networks.

This supplement introduces basic concepts and methods of queuing analysis that have wide applicability in manufacturing and service organizations. We focus only on simple models; other textbooks devoted exclusively to management science develop more complex models.

**ANALYZING QUEUES USING ANALYTICAL MODELS**

Many analytical queuing models exist, each based on unique assumptions about the nature of arrivals, service times, and other aspects of the system. Some of the common models are

1. Single- or multiple-channel with Poisson arrivals and exponential service times. (This is the most elementary situation.)

**Value-based queuing** is a method that allows organizations to prioritize customer calls based on their long-term value to the organization.
2. Single-channel with Poisson arrivals and arbitrary service times. (Service times may follow any probability distribution, and only the average and the standard deviation need to be known.)

3. Single-channel with Poisson arrivals and deterministic service times. (Service times are assumed to be constant.)

4. Single- or multiple-channel with Poisson arrivals, arbitrary service times, and no waiting line. (Waiting is not permitted. If the server is busy when a unit arrives, the unit must leave the system but may try to reenter at a later time.)

5. Single- or multiple-channel with Poisson arrivals, exponential service times, and a finite calling population. (A finite population of units is permitted to arrive for service.)

We illustrate the development of the basic queuing model for the problem of designing an automated check-in kiosk for passengers at an airport. Suppose that process design and facility-layout activities are currently being conducted for a new terminal at a major airport. One particular concern is the design and layout of the passenger check-in system. Most major airlines now use automated kiosks to speed up the process of obtaining a boarding pass with an electronic ticket. Passengers either enter a confirmation number or scan their electronic ticket to print a boarding pass. A queuing analysis of the system will help to determine if the system will provide adequate service to the airport passengers. To develop a queuing model, we must identify some important characteristics of the system: (1) the arrival distribution of the passengers, (2) the service-time distribution for the check-in operation, and (3) the waiting-line, or queue discipline for the passengers.

**Arrival Distribution**

Defining the arrival distribution for a waiting line consists of determining how many customers arrive for service in given periods of time, for example, the number of passengers arriving at the check-in kiosk during each 1-, 10-, or 60-minute period. Since the number of passengers arriving each minute is not a constant, we need to define a probability distribution that will describe the passenger arrivals. The choice of time period is arbitrary—as long as the same time period is used consistently—and is often determined based on the rate of arrivals and the ease by which the data can be collected. Generally, the slower the rate of arrivals, the longer the time period chosen.

For many waiting lines, the arrivals occurring in a given period of time appear to have a random pattern—that is, although we may have a good estimate of the total number of expected arrivals, each arrival is independent of other arrivals, and we cannot predict when it will occur. In such cases, a good description of the arrival pattern is obtained from the Poisson probability distribution:

\[
P(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{for } x = 0, 1, 2, \ldots
\]

where

- \(x\) = number of arrivals in a specific period of time
- \(\lambda\) = average, or expected, number of arrivals for the specific period of time
- \(e \approx 2.71828\)

For the passenger check-in process, the wide variety of flight schedules and the variation in passenger arrivals for the various flights cause the number of passengers arriving to vary substantially. For example, data collected from the actual operation of similar facilities show that in some instances, 20 to 25 passengers arrive during a 10-minute period. At other times, however, passenger arrivals drop to three or fewer passengers during a 10-minute period. Because passenger arrivals
cannot be controlled and appear to occur in an unpredictable fashion, a random arrival pattern appears to exist. Thus the Poisson probability distribution should provide a good description of the passenger-arrival pattern.

Airport planners have projected passenger volume through the year and estimate that passengers will arrive at an average rate of nine passengers per 10-minute period during the peak activity periods. Note that the choice of time period is arbitrary. We could have used an equivalent rate of 54 passengers per hour or 0.9 passengers per minute—as long as we are consistent in using the same time period in our analysis. Using the average, or mean arrival rate \( \lambda = 9 \), we can use the Poisson distribution defined in Equation (B.1) to compute the probability of \( x \) passenger arrivals in a 10-minute period.

\[
P(x) = \frac{9^x e^{-9}}{x!} \text{ for } x = 0, 1, 2, \ldots
\]

Sample calculations for \( x = 0, 5, \) and 10 passenger arrivals during a 1-minute period follow:

\[
P(0) = \frac{9^0 e^{-9}}{0!} = .0001
\]

\[
P(5) = \frac{9^5 e^{-9}}{5!} = .0607
\]

\[
P(10) = \frac{9^{10} e^{-9}}{10!} = .1186
\]

Using the Poisson probability distribution, we expect it to be very rare to have a 10-minute period in which no passengers \((x = 0)\) arrive for screening, since \(P(0) = .0001\). Five passenger arrivals occur with a probability \(P(5) = .0607\), and 10 with a probability of \(P(10) = .1186\). The probabilities for other numbers of passenger arrivals can also be computed. Exhibit B.1 shows the arrival distribution for passengers based on the Poisson distribution. In practice, you would want to record the actual number of arrivals per time period for several days or weeks and then compare the frequency distribution of the observed number of arrivals to the Poisson distribution to see if the Poisson distribution is a good approximation of the arrival distribution.

**Exhibit B.1**

Poisson Distribution of Passenger Arrivals
Service-Time Distribution

A service-time probability distribution is needed to describe how long it takes to check in a passenger at the kiosk. This length of time is referred to as the service time for the passenger. Although many passengers will complete the check-in process in a relatively short time, others might take a longer time because of unfamiliarity with the kiosk operation, ticketing problems, flight changes, and so on. Thus we expect service times to vary from passenger to passenger. In the development of waiting-line models, operations researchers have found that the exponential probability distribution can often be used to describe the service-time distribution. Equation (B.2) defines the exponential probability distribution

\[ f(t) = \mu e^{-\mu t} \quad \text{for } t \geq 0 \quad \text{(B.2)} \]

where

- \( t \) = service time (expressed in number of time periods)
- \( \mu \) = average or expected number of units that the service facility can handle in a specific period of time
- \( e \approx 2.71828 \)

It is important to use the same time period used for defining arrivals in defining the average service rate! If we use an exponential service-time distribution, the probability of a service being completed within \( t \) time periods is given by

\[ P(\text{Service time} \leq t \text{ Time periods}) = 1 - e^{-\mu t} \quad \text{(B.3)} \]

By collecting data on service times for similar check-in systems in operation at other airports, we find that the system can handle an average of 10 passengers per 10-minute period. Using a mean service rate of \( \mu = 10 \) customers per 10-minute period in Equation (B.3), we find that the probability of a check-in service being completed within \( t \) 10-minute periods is

\[ P(\text{Service time} \leq t \text{ 10-minute time periods}) = 1 - e^{-10t} \]

Now we can compute the probability that a passenger completes the service within any specified time, \( t \). For example, for 1 minute, we set \( t = 0.1 \) (as a fraction of a 10-minute period). Some example calculations are

\[ P(\text{Service time} \leq 1 \text{ minute}) = 1 - e^{-10(0.1)} = 1 - e^{-1} = .6321 \]
\[ P(\text{Service time} \leq 2.5 \text{ minutes}) = 1 - e^{-10(0.25)} = 1 - e^{-2.5} = .9179 \]

Thus, using the exponential distribution, we would expect 63.21 percent of the passengers to be serviced in 1 minute or less, and 91.79 percent in 2 1/2 minutes or less. Exhibit B.2 shows graphically the probability that \( t \) minutes or less will be required to service a passenger.

In the analysis of a specific waiting line, we want to collect data on actual service times to see if the exponential distribution assumption is appropriate. If you find other service-time patterns (such as a normal service-time probability distribution or a constant service-time distribution), the exponential distribution should not be used.

Queue Discipline

A queue discipline is the manner in which new arrivals are ordered or prioritized for service. For the airport problem, and in general for most customer-oriented waiting lines, the waiting units are ordered on a first-come, first-served (FCFS) basis—referred to as an FCFS queue discipline. Other types of queue disciplines are also prevalent. These include

- Shortest processing time (SPT), which we discussed in Chapter 14. SPT tries to maximize the number of units processed, but units with long processing times must wait long periods of time to be processed, if at all.
A random queue discipline provides service to units at random regardless of when they arrived for service. In some cultures, a random queue discipline is used for serving people instead of the FCFS rule.

Triage is used by hospital emergency rooms based on the criticality of the patient’s injury as patients arrive. That is, a patient with a broken neck receives top priority over another patient with a cut finger.

Preemption is the use of a criterion that allows new arrivals to displace members of the current queue and become the first to receive the service. This criterion could be wealth, society status, age, government position, and so on. Triage is a form of preemption based on the patient’s degree and severity of medical need.

Reservations and appointments allocate a specific amount of capacity at a specific time for a specific customer or processing unit. Legal and medical services, for example, book their day using appointment queuing disciplines.

A few of these queue disciplines are modeled analytically but most require simulation models to capture system queuing behavior. We will restrict our attention in this chapter to waiting lines with a FCFS queue discipline.

Queuing Behavior

People’s behavior in queues and service encounters is often unpredictable. Reneging is the process of a customer entering the waiting line but later deciding to leave the line and server system. Balking is the process of a customer evaluating the waiting line and server system and deciding not to enter the queue. In both situations, the customer leaves the system, may not return, and a current sale or all future sales may be lost. Most analytical models assume the customer’s behavior is patient and steady and they will not renege or balk, as such situations are difficult to model without simulation.

SINGLE-SERVER QUEUING MODEL

The queuing model presented in this section can be applied to waiting-line situations that meet these assumptions or conditions:

1. The waiting line has a single server.
2. The pattern of arrivals follows a Poisson probability distribution.
3. The service times follow an exponential probability distribution.
4. The queue discipline is first-come, first-served (FCFS).
5. No balking or reneging.
Because we have assumed that these conditions are applicable to the airport check-in problem, we can use this queuing model to analyze the operation. We have already concluded that the mean arrival rate is $\lambda = 9$ passengers per 10-minute period and that the mean service rate is $\mu = 10$ passengers per 10-minute period. Using the assumptions of Poisson arrivals and exponential service times, quantitative analysts have developed the following expressions to define the operating characteristics of a single-channel waiting line:

1. the probability that the service facility is idle (that is, the probability of 0 units in the system):
   \[ P_0 = (1 - \lambda/\mu) \]  \hspace{1cm} (B.4)

2. the probability of $n$ units in the system:
   \[ P_n = (\lambda/\mu)^n P_0 \]  \hspace{1cm} (B.5)

3. the average number of units waiting for service:
   \[ L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} \]  \hspace{1cm} (B.6)

4. the average number of units in the system:
   \[ L = L_q + \frac{\lambda}{\mu} \]  \hspace{1cm} (B.7)

5. the average time a unit spends waiting for service:
   \[ W_q = L_q/\lambda \]  \hspace{1cm} (B.8)

6. the average time a unit spends in the system (waiting time plus service time):
   \[ W = W_q + 1/\mu \]  \hspace{1cm} (B.9)

7. the probability that an arriving unit has to wait for service:
   \[ P_w = \frac{\lambda}{\mu} \]  \hspace{1cm} (B.10)

The values of the mean arrival rate, $\lambda$, and the mean service rate, $\mu$, are clearly important components in these formulas. From Equation (B.10), we see that the ratio of these two values, $\lambda/\mu$, is simply the probability that an arriving unit must wait because the server is busy. Thus, $\lambda/\mu$ is often referred to as the utilization factor for the waiting line. The formulas for determining the operating characteristics of a single-server waiting line presented in Equations (B.4) through (B.10) are applicable only when the utilization factor, $\lambda/\mu$, is less than 1. This condition occurs when the mean service rate, $\mu$, is greater than the mean arrival rate, $\lambda$, and hence when the service rate is sufficient to process or service all arrivals.

Returning to the airport check-in problem, we see that with $\lambda = 9$ and $\mu = 10$, we can use Equations (B.4) through (B.10) to determine the operating characteristics of the screening operation. This is done as follows:

\[ P_0 = (1 - \lambda/\mu) = (1 - 9/10) = .10 \]

\[ L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{9^2}{10(10 - 9)} = 81/10 = 8.1 \text{ passengers} \]

\[ L = L_q + \frac{\lambda}{\mu} = 8.1 + 9/10 = 9.0 \text{ passengers} \]

\[ W_q = L_q/\lambda = 8.1/9 = .9 \text{ (Note that this refers to the number of 10-minute periods, or equivalently, 9 minutes per passenger)} \]

\[ W = W_q + 1/\mu = 0.9 \text{ hour} + 1/10 \text{ hour} = \text{one 10-minute period, or equivalently, 10 minutes per passenger} \]

\[ P_w = \frac{\lambda}{\mu} = 9/10 = .90 \]
Using this information, we can learn several important things about the check-in operation. In particular, we see that passengers wait an average of 9 minutes at the kiosk. With this as the average, many passengers wait even longer. In airport operations with passengers rushing to meet plane connections, this waiting time might be judged to be undesirably high. In addition, the facts that the average number of passengers waiting in line is 8.1 and that 90 percent of the arriving passengers must wait to check in might suggest to the operations manager that something should be done to improve the efficiency of the process.

These operating characteristics are based on the assumption of an arrival rate of 9 and a service rate of 10 per 10-minute period. As the figures are based on airport planners’ estimates, they are subject to forecasting errors. It is easy to examine the effects of a variety of assumptions about arrival and service rates on the operating characteristics by using a spreadsheet such as the one in Exhibit B.3. In this spreadsheet, we compute the operating characteristics for the baseline assumptions and also examine the effect of changes in the mean arrival rate from 7 to 10 passengers per period.

What do the data in this figure tell us about the design of this process? They tell us that if the mean arrival rate is 7 passengers per period, the system functions acceptably. On the average, only 1.63 passengers are waiting and the average waiting time of 0.23(10 minutes) = 2.3 minutes appears acceptable. However, we see that the mean arrival rate of 9 passengers per period provides undesirable waiting characteristics, and if the rate increases to 10 passengers per period, the system as proposed is completely inadequate. When $\lambda = \mu$, the operating characteristics are not defined, meaning that these times and numbers of passengers grow infinitely large (that is, when $\lambda = \mu \rightarrow \infty$). These results show that airport planners need to consider design modifications that will improve the efficiency of the check-in process.

If a new process can be designed that will improve the passenger-service rate, Equations (B.4) through (B.10) can be used to predict operating characteristics under any revised mean service rate, $\mu$. Developing a spreadsheet with alternative mean service rates provides the information to determine which, if any, of the screening facility designs can handle the passenger volume acceptably.

Computing the probability of more or less than $x$ units arriving requires us to use Equation (B.5) and the following two equations:

$$P(\text{Number of arrivals} > x) = 1 - P(\text{Number of arrivals} \leq x) \quad (B.11)$$

$$P(\text{Number of arrivals} < x) = 1 - P(\text{Number of arrivals} \geq x) \quad (B.12)$$

---

**Exhibit B.3**

Spreadsheet for Single-Server Queuing Model

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Single-Server Queuing Model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Lambda</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>Mu</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Probability system is empty</td>
<td>0.30</td>
<td>0.20</td>
<td>0.10</td>
</tr>
<tr>
<td>7</td>
<td>Average number in queue</td>
<td>1.63</td>
<td>3.20</td>
<td>8.10</td>
</tr>
<tr>
<td>8</td>
<td>Average number in system</td>
<td>2.33</td>
<td>4.00</td>
<td>9.00</td>
</tr>
<tr>
<td>9</td>
<td>Average time in queue</td>
<td>0.23</td>
<td>0.40</td>
<td>0.90</td>
</tr>
<tr>
<td>10</td>
<td>Average waiting time in system</td>
<td>0.33</td>
<td>0.50</td>
<td>1.00</td>
</tr>
<tr>
<td>11</td>
<td>Probability arrival has to wait</td>
<td>0.70</td>
<td>0.80</td>
<td>0.90</td>
</tr>
</tbody>
</table>
These equations are used to simplify the calculations. For example, to find the probability that more than 4 customers are waiting for service, we would need to sum the probabilities associated with 5, 6, 7, . . . up to a (theoretically) infinite number. From Equation (B.11), we simply need to find the probabilities associated with 4 or less, sum them up, and subtract them from 1.0. With \( \lambda = 9 \) and \( \mu = 10 \) and using Equation (B.5), we have

\[
\begin{align*}
P_0 &= (1 - \lambda/\mu) = (1 - 9/10) = .1000 \\
P_1 &= (\lambda/\mu)P_0 = (9/10)^1(.1) = .0900 \\
P_2 &= (\lambda/\mu)P_0 = (9/10)^2(.1) = .0810 \\
P_3 &= (\lambda/\mu)P_0 = (9/10)^3(.1) = .0729 \\
P_4 &= (\lambda/\mu)P_0 = (9/10)^4(.1) = .0656 \\
P(\text{Number of arrivals} \geq 4) &= .0656 \\
\end{align*}
\]

Therefore, the probability that more than 4 customers would be waiting for service is \( 1 - .4095 = .5905 \).

**MULTIPLE-SERVER QUEUING MODEL**

A logical extension of a single-server waiting line is to have multiple servers, similar to those you are familiar with at many banks. By having more than one server, the check-in process can be dramatically improved. In this situation, customers wait in a single line and move to the next available server. Note that this is a different situation from one in which each server has a distinct queue, such as with highway tollbooths, bank teller windows, or supermarket checkout lines. In such situations, customers might “jockey” for position between servers (channels). **Jockeying** is the process of customers leaving one waiting line to join another in a multiple-server (channel) configuration. The model we present assumes that all servers are fed from a single waiting line. Exhibit B.4 is a diagram of this system.

In this section we present formulas that can be used to compute various operating characteristics for a multiple-server waiting line. The model we will use can be applied to situations that meet these assumptions:

1. The waiting line has two or more identical servers.

Exhibit B.4
A Two-Server Queuing System
2. The arrivals follow a Poisson probability distribution with a mean arrival rate of \( \lambda \).
3. The service times have an exponential distribution.
4. The mean service rate, \( \mu \), is the same for each server.
5. The arrivals wait in a single line and then move to the first open server for service.
6. The queue discipline is first-come, first-served (FCFS).
7. No balking or reneging is allowed.

Using these assumptions, operations researchers have developed formulas for determining the operating characteristics of the multiple-server waiting line. Let

- \( k \) = number of channels
- \( \lambda \) = mean arrival rate for the system
- \( \mu \) = mean service rate for each channel

The following equations apply to multiple-server waiting lines for which the overall mean service rate, \( k\mu \), is greater than the mean arrival rate, \( \lambda \). In such cases, the service rate is sufficient to process all arrivals.

1. the probability that all \( k \) service channels are idle (that is, the probability of zero units in the system):
   \[
P_0 = \frac{1}{\sum_{n=0}^{k-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^k \mu}{(k-1)! k \mu - \lambda}} \tag{B.13}
   \]
2. the probability of \( n \) units in the system:
   \[
P_n = \frac{(\lambda/\mu)^n}{k! k^{n-k} P_0} \quad \text{for } n > k
   \]
   \[
P_n = \frac{(\lambda/\mu)^n}{n!} P_0 \quad \text{for } 0 \leq n \leq k \tag{B.14}
   \]
3. the average number of units waiting for service:
   \[
   L_q = \frac{(\lambda/\mu)^k \lambda \mu}{(k-1)!(k\mu - \lambda)^2 P_0} \tag{B.15}
   \]
4. the average number of units in the system:
   \[
   L = L_q + \lambda/\mu \tag{B.16}
   \]
5. the average time a unit spends waiting for service:
   \[
   W_q = L_q/\lambda \tag{B.17}
   \]
6. the average time a unit spends in the system (waiting time plus service time):
   \[
   W = W_q + 1/\mu \tag{B.18}
   \]
7. the probability that an arriving unit must wait for service:
   \[
   P_w = \frac{1}{k! \left( \frac{\lambda}{\mu} \right)^k} \frac{k \mu}{k \mu - \lambda} P_0 \tag{B.19}
   \]

Although the equations describing the operating characteristics of a multipleserver queuing model with Poisson arrivals and exponential service times are somewhat more complex than the single-server equations, they provide the same information and are used exactly as we used the results from the single-channel model. To simplify the use of Equations (B.13) through (B.19), Exhibit B.5 shows values of \( P_0 \) for selected values of \( \lambda/\mu \). Note that the values provided correspond to cases for which \( k\mu > \lambda \); hence the service rate is sufficient to service all arrivals.
For an application of the multiple-server waiting-line model, we return to the airport check-in problem and consider the desirability of expanding the screening facility to provide two kiosks. How does this design compare to the single-server alternative?

We answer this question by applying Equations (B.13) through (B.19) for \( k = 2 \) servers. Using an arrival rate of \( \lambda = 9 \) passengers per period and \( \mu = 10 \) passengers per period for each of the kiosks, we have these operating characteristics:

\[
P_0 = 0.3793 \quad \text{(from Exhibit B.5 for } \lambda/\mu = 0.9 \text{ and } k = 2)\]

\[
L_q = \frac{(9/10)^2(9)(10)}{(2 - 1)!(20 - 9)^2} (0.3793) = 0.23 \text{ passengers}\]

\[
L = 0.23 + \frac{9}{10} = 1.13 \text{ passengers}\]

\[
W_q = \frac{0.23}{9} = 0.026 \quad \text{multiples of 10-minute periods, or 0.26 minutes/passerger}\]

For an application of the multiple-server waiting-line model, we return to the airport check-in problem and consider the desirability of expanding the screening facility to provide two kiosks. How does this design compare to the single-server alternative?

We answer this question by applying Equations (B.13) through (B.19) for \( k = 2 \) servers. Using an arrival rate of \( \lambda = 9 \) passengers per period and \( \mu = 10 \) passengers per period for each of the kiosks, we have these operating characteristics.

\[
P_0 = \frac{(9/10)^2(9)(10)}{(2 - 1)!(20 - 9)^2} (0.3793) = 0.23 \text{ passengers}\]

\[
L = 0.23 + \frac{9}{10} = 1.13 \text{ passengers}\]

\[
W_q = \frac{0.23}{9} = 0.026 \quad \text{multiples of 10-minute periods, or 0.26 minutes/passerger}\]
These operating characteristics suggest that the two-server operation would handle the volume of passengers extremely well. Specifically, note that the total time in the system is an average of only 1.26 minutes per passenger, which is excellent. The percentage waiting is 27.9 percent, which is acceptable, especially in light of the short average waiting time.

Exhibit B.6 is a spreadsheet designed to compute operating characteristics for up to eight servers in the multiple-server queuing model using the arrival and service rates for the security-screening example. (The Microsoft Excel function HLOOKUP is used to compute the summation and the term $(\lambda/\mu)^n/n!$ in the denominator of $P_0$. The function FACT computes factorials.) With three servers, we see a significant improvement over two servers in the operating characteristics; beyond this, the improvement is negligible. In addition, we can use the spreadsheet to show that even if the mean arrival rate for passengers exceeds the estimated 9 passengers per hour, the two-channel system should operate nicely.

What is the probability that less than four customers are waiting for service when $k = 2$, $\lambda = 9$, and $\mu = 10$ passengers per time period? Using the spreadsheet, we can calculate the probabilities when $x = 0$ as .3793, $x = 1$ as .3414, $x = 2$ as .1536, $x = 3$ as .0691, and $x = 4$ as .0311. Using Equation (B.13), we sum the probabilities from $P_0$ to $P_4$ and then subtract from 1 to arrive at the probability of four or more customers waiting for service at $1 - .9745 = .0255$. Adding the second server greatly improves system performance, as the results show.
THE ECONOMICS OF WAITING-LINE ANALYSIS

As we have shown, queuing models can be used to determine operating performance of a waiting-line system. In the economic analysis of waiting lines, we seek to use the information provided by the queuing model to develop a cost model for the waiting line under study. Then we can use the model to help the manager balance the cost of customers having to wait for service against the cost of providing the service. This is a vital issue for all operations managers (see the OM Spotlight on airport security screening).

In developing a cost model for the check-in problem, we will consider the cost of passenger time, both waiting time and servicing time, and the cost of operating the system. Let $C_W = \text{the waiting cost per hour per passenger}$ and $C_S = \text{the hourly cost associated with each server}$. Clearly, the passenger waiting-time cost cannot be accurately determined; managers must estimate a reasonable value that might reflect the potential loss of future revenue should a passenger switch to another airport or airline because of perceived unreasonable delays. This is called the *imputed cost of waiting*. Suppose $C_W$ is estimated to be $50 per hour, or $0.83 per minute. The cost of operating each service facility is more easily determined, as it consists of the wages of any personnel and the cost of equipment, including maintenance. For automated systems, this is usually quite small. Let us assume that $C_S = $10 per hour, or $0.167 per minute. Therefore, the total cost per minute is $C_W L + C_S k$, where $L = \text{average number of passengers in the system}$ and $k = \text{number of servers}$. Exhibit B.7 summarizes the cost for the one- and two-server scenarios. We clearly see the economic advantages of a two-server system.

### Exhibit B.7
Economic Analysis of Check-In System Design

<table>
<thead>
<tr>
<th>System</th>
<th>$k$</th>
<th>System Cost</th>
<th>$L$</th>
<th>Passenger Cost</th>
<th>Total Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-server</td>
<td>1</td>
<td>0.167(1) = 0.167</td>
<td>9</td>
<td>0.83(9) = 7.47</td>
<td>$7.64</td>
</tr>
<tr>
<td>Two-server</td>
<td>2</td>
<td>0.167(2) = 0.334</td>
<td>1.13</td>
<td>0.83(1.13) = 0.94</td>
<td>$1.27</td>
</tr>
</tbody>
</table>

OM SPOTLIGHT

**Airport Security Wait Times**

The U.S. Transportation Security Administration (TSA) set up a new web site that provides airport-by-airport information on the average wait times. The site provides hourly and daily average wait times based on last month’s data. A longer-term goal is to provide real-time hourly updates. In 2004, the longest (maximum) time waiting in line to get to the metal detector was 36 minutes at a major U.S. airport. Average waiting times range from a few minutes to 30 minutes. For example, the TSA recorded the wait at the main security checkpoint at Hartsfield International Airport in Atlanta at 7 A.M. on Monday, August 9, 2004 averaged 26 minutes. Monday morning is a peak time for most airports. The data are collected by security screeners who give passengers a card with their arrival time on it, which is collected when they get to the metal detector.
THE PSYCHOLOGY OF WAITING

Customers become frustrated when a person enters a line next to them and receives service first. Of course, that customer feels a certain sense of satisfaction. People expect to be treated fairly; in queuing situations that usually means “first-come, first-served.” In the mid-1960s, Chemical Bank was one of the first to switch to a serpentine line (one line feeding into several servers) from multiple parallel lines. American Airlines copied this at its airport counters and most others followed suit. Studies have shown that customers are happier when they wait in a serpentine line, rather than in parallel lines, even if that type of line increases their wait.

Understanding the psychological perception of waiting is as important in addressing queuing problems as are analytical approaches. Creative solutions that do not rely on technical approaches can be quite effective. One example involved complaints of tenants waiting for elevators in a high-rise building. Rather than pursuing an expensive technical solution of installing a faster elevator, the building manager installed mirrors in the elevator lobbies to help the tenants pass the time. This is commonly found in many hotels today. In other elevator lobbies, we often see art or restaurant menus to distract patrons. Another example occurred at the Houston airport. Passengers complained about long waits when picking up their baggage. The airline solved the problem by moving the baggage to the farthest carousel from the planes. While the total time to deliver the baggage was not changed, the fact that passengers had to walk farther and wait less eliminated the complaints.

Nothing is worse than not knowing when the next bus will arrive. Not knowing how long a wait will be creates anxiety. To alleviate this kind of uncertainty, the Disney theme parks inform people how long a wait to expect by placing signs at various points along the queue. Chemical Bank pays $5 to customers who wait in line more than 7 minutes. This interval was chosen because research indicated that waits up to 10 minutes were tolerable. Customers have provided good feedback; they do not seem to mind waiting longer if they receive something for it.

Florida Power and Light developed a system that informed customers of the estimated waiting time for telephone calls, allowing customers to call back later if the wait would be too long. Consumer research revealed that customers would wait 94 seconds without knowing the length of wait. It also showed that customers began to be dissatisfied after waiting about 2 minutes. But when customers knew the length of wait, they were willing to wait 105 seconds longer—a total of 199 seconds! Thus, Florida Power and Light knew that it could buy more time, without sacrificing customer satisfaction, by giving customers a choice of holding for a predicted period of time or deferring the call to a later time. The system, called Smartqueue, was implemented, and virtually all customers considered it helpful in subsequent satisfaction surveys. From the company’s perspective, Smartqueue increased the time customers were willing to wait without being dissatisfied by an appreciable amount.

Other methods of changing customers’ perceptions involve distractions. Time spent without anything to do seems longer than occupied time. Airlines and rental car firms divide their processes into stages to make the process seem shorter with breaks in service for both the service provider’s and customer’s benefit. Hospitals try to reduce the perception of waiting all day in the hospital by separating patient parking, admission, blood test, x-rays, examination, and so on from one another. Guests waiting for a ride at Disney World seldom see the entire queue, which can have hundreds of people. Amusement parks might also have roving entertainers to distract the waiting crowds. As early as 1959, the Manhattan Savings Bank offered live entertainment and even dog and boat shows during the busy lunchtime hours.
Supermarkets place “impulse” items such as candy, batteries, and other small items as well as magazines near checkouts to grab customers’ attention. The Postal Service has been experimenting with video displays that not only distract customers but also inform them of postal procedures so that they can speed up their transactions.

Technology is alleviating queuing in many service industries today. For example, rental car firms use automatic tellers for fast check in and out and are working on radio frequency technology to entirely skip waiting in lines to get or return a vehicle. Airlines allow their customers to print out boarding passes at airport kiosks or on their own printer to speed the check-in process. Thus, queuing in operations management entails much more than some analytical calculations and requires good management skills.

**SOLVED PROBLEMS**

**SOLVED PROBLEM #1**

The reference desk of a large library receives requests for assistance at a mean rate of 10 requests per hour, and it is assumed that the desk has a mean service rate of 12 requests per hour.

a. What is the probability that the reference desk is idle?

b. What is the average number of requests that will be waiting for service?

c. What is the average number of requests in the system?

d. What is the average waiting time plus service time for a request for assistance?

e. What is the utilization factor?

f. What is the probability of more than three requests?

**Solution:**

a. \[ P_0 = (1 - \lambda/\mu) = (1 - 10/12) = .1667 \]  
   (Equation B.4)

b. \[ L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{10^2}{12(12 - 10)} = 4.1667 \text{ requests} \]  
   (Equation B.6)

c. \[ L = L_q + \lambda/\mu = 4.1667 + 10/12 = 5.000 \text{ requests} \]  
   (Equation B.7)

d. \[ W_q = L_q/\lambda = 4.1667/10 = 0.4167 \text{ hour} \]  
   (Equation B.8)

\[ W = W_q + 1/\mu = 0.4167 + 1/12 = 0.5 \text{ hour, or 30 minutes} \]  
   (Equation B.9)

e. \[ P_\infty = \lambda/\mu = 10/12 = .8333 \]  
   (Equation B.10)

f. Use Equation (B.5) to compute the following:

\[ P_0 = (1 - \lambda/\mu) = (1 - 10/12) = .1667 \]

\[ P_1 = (\lambda/\mu)^n P_0 = (10/12)^1(.1667) = .1389 \]

\[ P_2 = (\lambda/\mu)^n P_0 = (10/12)^2(.1667) = .1157 \]

\[ P_3 = (\lambda/\mu)^n P_0 = (10/12)^3(.1667) = .0965 \]

Sum of \( P_i \leq x \) = .5178

Using Equation (B.13), we sum the probabilities from \( P_0 \) to \( P_3 \) and then subtract from 1 to arrive at the probability of more than three requests waiting for service of \( 1 - .5178 = .4822 \).

**SOLVED PROBLEM #2**

A fast-food franchise operates a drive-up window. Orders are placed at an intercom station at the back of the parking lot. After placing an order, the customer pulls up and waits in line at the drive-up window until the cars in front have been served. By hiring a second person to help take and fill orders, the manager hopes to improve service. With one person filling orders, the average service time for a drive-up customer is 2 minutes;
with a second person working, the average service time can be reduced to 1 minute, 15 seconds. Note that the drive-up window operation with two people is still a single-channel waiting line. However, with the addition of the second person, the average service time can be decreased. Cars arrive at the rate of 24 per hour.

a. Determine the average waiting time when one person is working the drive-up window.

b. With one person working the drive-up window, what percentage of time will that person not be occupied serving customers?

c. Determine the average waiting time when two people are working at the drive-up window.

d. With two persons working the drive-up window, what percentage of time will no one be occupied serving drive-up customers?

e. Would you recommend hiring a second person to work the drive-up window? Justify your answer.

Solution:

\[ \lambda = 24, \ \mu = \frac{60}{2} = 30 \]

\[ L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{24^2}{30(30 - 24)} = 3.2 \]

\[ W_q = \frac{L_q}{\lambda} = 0.1333 \text{ hour (8 minutes)} \]

\[ P_0 = 1 - \frac{\lambda}{\mu} = 1 - \frac{24}{30} = .20 \]

\[ \lambda = 24, \ \mu = \frac{60}{1.25} = 48 \]

\[ L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{24^2}{48(48 - 24)} = 0.5 \]

\[ W_q = \frac{L_q}{\lambda} = 0.0208 \text{ hour (1.25 minutes)} \]

\[ P_0 = 1 - \frac{\lambda}{\mu} = 1 - \frac{24}{48} = 0.50 \]

e. Yes, waiting time is reduced from \( W_q = 8 \) to \( W_q = 1.25 \).

KEY TERMS AND CONCEPTS

- Balking
- Economics of waiting
- Exponential probability distribution
- Jockeying
- Mean arrival rate
- Mean service rate
- Multiple-server models
- Poisson probability distribution
- Psychology of waiting
- Queue discipline (queue priority rule)
- Queue/waiting line
- Queuing performance measures
- Reneging
- Single-server models
- Types of queue disciplines
- Utilization factor
- Value-based queuing

QUESTIONS FOR REVIEW AND DISCUSSION

1. What are six typical performance measures for waiting-line (queuing) models?
2. What information is necessary to analyze a waiting line?
3. What probability distributions are used in basic queuing models? Explain.
4. Define a “queue discipline” and why this is important?
5. Explain how resource utilization is measured in single- and multiple-server queuing models.
7. List the assumptions underlying the basic single-channel waiting-line model.
8. What are the assumptions of the multiple-channel waiting-line model?
9. Can queuing models predict the size of the queue (waiting line) when vehicle traffic is limited to two lanes instead of three lanes? What if traffic is reduced to one lane? Make up a simple problem and explain what you discover.

10. Does customers’ value over their buying life figure into basic queuing model logic? Explain.
11. How do restaurants handle the psychology of waiting as customers arrive for dinner? Explain.
12. Explain a waiting-line multiple-server situation where you jockeyed for position and it improved or hindered your service. Did other customers get mad? Explain.

**PROBLEMS AND ACTIVITIES**

The following waiting-line problems are all based on the assumptions of Poisson arrivals and exponential service times.

1. Arrivals to a single-server queue occur at an average rate of 3 per minute. Develop the probability distribution for the number of arrivals for $x = 0$ through 10.

2. Trucks using a single-server loading dock have a mean arrival rate of 12 per day. The loading/unloading rate is 18 per day.
   a. What is the probability that the truck dock will be idle?
   b. What is the average number of trucks waiting for service?
   c. What is the average time a truck waits for the loading or unloading service?
   d. What is the probability that a new arrival will have to wait?
   e. What is the probability that more than three trucks are waiting for service?

3. A mail-order nursery specializes in European beech trees. New orders, which are processed by a single shipping clerk, have a mean arrival rate of six per day and a mean service rate of eight per day.
   a. What is the average time an order spends in the queue waiting for the clerk to begin service?
   b. What is the average time an order spends in the system?

4. Assume trucks arriving for loading/unloading at a truck dock form a single-server waiting line. The mean arrival rate is four trucks per hour and the mean service rate is five trucks per hour.
   a. What is the probability that the truck dock will be idle?
   b. What is the average number of trucks in the queue?
   c. What is the average number of trucks in the system?
   d. What is the average time a truck spends in the queue waiting for service?
   e. What is the average time a truck spends in the system?
   f. What is the probability that an arriving truck will have to wait?
   g. What is the probability that more than two trucks are waiting for service?

5. Marty’s Barber Shop has one barber. Customers arrive at a rate of 2.2 per hour, and haircuts are given at an average rate of 5 customers per hour.
   a. What is the probability that the barber is idle?
   b. What is the probability that one customer is receiving a haircut and no one is waiting?
   c. What is the probability that one customer is receiving a haircut and one customer is waiting?
   d. What is the probability that one customer is receiving a haircut and two customers are waiting?
   e. What is the probability that more than two customers are waiting?
   f. What is the average time a customer waits for service?

6. Trosper Tire Company has decided to hire a new mechanic to handle all tire changes for customers ordering new tires. Two mechanics are available for the job. One mechanic has limited experience and can be hired for $7 per hour. It is expected that this mechanic can service an average of three customers per hour. A mechanic with several years of experience is also being considered for the job. This
mechanic can service an average of four customers per hour, but must be paid $10 per hour. Assume that customers arrive at the Trosper garage at the rate of two per hour.

a. Compute waiting-line operating characteristics for each mechanic.
b. If the company assigns a customer-waiting cost of $15 per hour, which mechanic provides the lower operating cost?

7. Agan Interior Design provides home and office decorating assistance. In normal operation an average of 2.5 customers arrive per hour. One design consultant is available to answer customer questions and make product recommendations. The consultant averages 10 minutes with each customer.

a. Compute operating characteristics for the customer waiting line.
b. Service goals dictate that an arriving customer should not wait for service more than an average of 5 minutes. Is this goal being met? What action do you recommend?
c. If the consultant can reduce the average time spent with customers to 8 minutes, will the service goal be met?

8. Pete’s Market is a small local grocery store with one checkout counter. Shoppers arrive at the checkout lane at an average rate of 15 customers per hour and the average order takes 3 minutes to ring up and bag. What information would you develop to help Pete analyze the current operation? If Pete does not want the average time waiting for service to exceed 5 minutes, what would you tell him about the current system?

9. Refer to Problem 8. After reviewing the analysis, Pete felt it would be desirable to hire a full-time person to assist in the checkout operation. Pete believed that if this employee assisted the cashier, average service time could be reduced to 2 minutes. However, Pete was also considering installing a second checkout lane, which could be operated by the new person. This would provide a two-server system with the average service time of 3 minutes for each customer. Should Pete use the new employee to assist on the current checkout counter or to operate a second counter? Justify your recommendation.

10. Keuka Park Savings and Loan currently has one drive-in teller window. Cars arrive at a mean rate of 10 per hour. The mean service rate is 12 cars per hour.

a. What is the probability that the service facility will be idle?
b. If you were to drive up to the facility, how many cars would you expect to see waiting and being serviced?
c. What is the average time waiting for service?
d. What is the probability an arriving car will have to wait?
e. What is the probability that more than four vehicles are waiting for service?
f. As a potential customer of the system, would you be satisfied with these waiting-line characteristics? How do you think managers could go about assessing its customers’ feelings about the current system?

11. To improve its customer service, Keuka Park Savings and Loan (Problem 10) wants to investigate the effect of a second drive-in teller window. Assume a mean arrival rate of 10 cars per hour. In addition, assume a mean service rate of 12 cars per hour for each window. What effect would adding a new teller window have on the system? Does this system appear acceptable?

12. Fore and Aft Marina is a new marina planned for a location on the Ohio River near Madison, Indiana. Assume that Fore and Aft decides to build one docking facility and expects a mean arrival rate of 5 boats per hour and a mean service rate of 10 boats per hour.

a. What is the probability that the boat dock will be idle?
b. What is the average number of boats that will be waiting for service?
c. What is the average time a boat will wait for service?
d. What is the average time a boat will spend at the dock?
e. What is the probability that more than 2 boats are waiting for service?
f. If you were the owner, would you be satisfied with this level of service?

13. The owner of the Fore and Aft Marina in Problem 12 is investigating the possibility of adding a second dock. Assume a mean arrival rate of 5 boats per hour for the marina and a mean service rate of 10 boats per hour for each server.

a. What is the probability that the boat dock will be idle?
b. What is the average number of boats that will be waiting for service?
c. What is the average time a boat will wait for service?
d. What is the average time a boat will spend at the dock?
e. If you were the owner, would you be satisfied with this level of service?

14. The City Beverage Drive-Thru is considering a two-server system. Cars arrive at the store at the mean rate of 6 per hour. The service rate for each server is 10 per hour.
   a. What is the probability that both servers are idle?
   b. What is the average number of cars waiting for service?
   c. What is the average time waiting for service?
   d. What is the average time in the system?
   e. What is the probability of having to wait for service?

15. Consider a two-server waiting line with a mean arrival rate of 50 per hour and a mean service rate of 75 per hour for each server.
   a. What is the probability that both servers are idle?
   b. What is the average number of cars waiting for service?
   c. What is the average time waiting for service?
   d. What is the average time in the system?
   e. What is the probability of having to wait for service?

16. For a two-server waiting line with a mean arrival rate of 14 per hour and a mean service rate of 10 per hour per server, determine the probability that an arrival must wait. What is the probability of waiting if the system is expanded to three servers?

17. Big Al’s Quickie Carwash has two wash bays. Each bay can wash 15 cars per hour. Cars arrive at the carwash at the rate of 15 cars per hour on the average, join the waiting line, and move to the next open bay when it becomes available.
   a. What is the average time waiting for a bay?
   b. What is the probability that a customer will have to wait?
   c. As a customer of Big Al’s, do you think the system favors the customer? If you were Al, what would be your attitude toward this service level?

18. Refer to the Agan Interior Design situation in Problem 7. Agan is evaluating two alternatives:
   1. use one consultant with an average service time of 8 minutes per customer;
   2. expand to two consultants, each of whom has an average service time of 10 minutes per customer.

If the consultants are paid $16 per hour and the customer waiting time is valued at $25 per hour, should Agan expand to the two-consultant system? Explain.

19. Refer to Solved Problem 2. Space is available to install a second drive-up window adjacent to the first. The manager is considering adding such a window. One person will be assigned to service customers at each window.
   a. Determine the average customer waiting time for this two-server system.
   b. What percentage of the time will both windows be idle?
   c. What design would you recommend: one attendant at one window, two attendants at one window, or two attendants and two windows with one attendant at each window?

20. Design a spreadsheet similar to Exhibit B.3 to study changes in the mean service rate from 10 to 15 for $\lambda = 9$ passengers per minute.

21. Using the spreadsheet in Exhibit B.6 (Multiple-Server Queue.xls), determine the effect of increasing passenger arrival rates of 10, 12, 14, 16, and 18 on the operating characteristics of the airport security screening example.

### CASES

#### BOURBON COUNTY COURT

“Why don’t they buy another copying machine for this office? I waste a lot of valuable time fooling with this machine when I could be preparing my legal cases,” noted H. C. Morris, as he waited in line. The self-service copying machine was located in a small room immediately outside the entrance of the courtroom. Morris was the county attorney. He often copied his own papers, as did other lawyers, to keep his legal cases and work confidential. This protected the privacy of his clients as well as his professional and personal ideas about the cases.

He also felt awkward at times standing in line with secretaries, clerks of the court, other attorneys, police
officers and sheriffs, building permit inspectors, and the dog warden—all trying, he thought, to see what he was copying. The line for the copying machine often extended out into the hallways of the courthouse.

Morris mentioned his frustration with the copying machine problem to Judge Hamlet and his summer intern, Dot Gifford. Gifford was home for the summer and working toward a joint MBA/JD degree from a leading university.

“Mr. Morris, there are ways to find out if that one copying machine is adequate to handle the demand. If you can get the judge to let me analyze the situation, I think I can help out. We had a similar problem at the law school with word processors and at the business school with student lab microcomputers.”

The next week Judge Hamlet gave Gifford the go-ahead to work on the copying machine problem. He asked her to write a management report on the problem with recommendations so he could take it to the Bourbon County Board of Supervisors for their approval. The board faced deficit spending last fiscal year, so the trade-offs between service and cost must be clearly presented to the board.

Gifford’s experience with analyzing similar problems at school helped her know what type of information and data were needed. After several weeks of working on this project, she developed the information contained in Exhibits B.8, B.9, and B.10.

Gifford was not quite as confident in evaluating this situation as others because the customer mix and associated labor costs seemed more uncertain in the county courthouse. In the law school situation, only secretaries used the word processing terminals; in the business school situation, students were the ones complaining about long waiting times to get on a microcomputer terminal. Moreover, the professor guiding these two

---

Exhibit B.8
Bourbon County Court—
Customer Arrivals per Hour*

<table>
<thead>
<tr>
<th>Customer Arrivals in One Hour</th>
<th>Customer Arrivals in One Hour</th>
<th>Customer Arrivals in One Hour</th>
<th>Customer Arrivals in One Hour</th>
<th>Customer Arrivals in One Hour</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 5 11 10 21 3 31 11 41 14</td>
<td>2 9 12 17 22 9 32 8 42 7</td>
<td>3 7 13 18 23 11 33 9 43 4</td>
<td>4 13 14 14 24 10 34 8 44 7</td>
<td>5 7 15 11 25 12 35 6 45 7</td>
</tr>
<tr>
<td>6 7 16 16 26 4 36 8 46 2</td>
<td>7 7 17 5 27 8 37 14 47 4</td>
<td>8 11 18 6 28 9 38 12 48 7</td>
<td>9 8 19 8 29 9 39 11 49 2</td>
<td>10 6 20 13 30 9 40 15 50 8</td>
</tr>
</tbody>
</table>

*A sample of customer arrivals at the copying machine was taken for five consecutive 9-hour work days plus 5 hours on Saturday for a total of 50 observations. The mean arrival rate is 8.92 arrivals per hour.

Exhibit B.9
Bourbon County Court—
Copying Service Times*

<table>
<thead>
<tr>
<th>Obs. No.</th>
<th>Hours per Job</th>
<th>Obs. No.</th>
<th>Hours per Job</th>
<th>Obs. No.</th>
<th>Hours per Job</th>
<th>Obs. No.</th>
<th>Hours per Job</th>
<th>Obs. No.</th>
<th>Hours per Job</th>
<th>Obs. No.</th>
<th>Hours per Job</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0.0700</td>
<td>11 0.1253</td>
<td>21 0.1754</td>
<td>31 0.0752</td>
<td>41 0.2005</td>
<td>2 0.1253</td>
<td>12 0.1754</td>
<td>22 0.0700</td>
<td>32 0.1002</td>
<td>42 0.0501</td>
<td>3 0.0752</td>
<td>13 0.0301</td>
</tr>
<tr>
<td>4 0.2508</td>
<td>14 0.1002</td>
<td>24 0.0752</td>
<td>34 0.0752</td>
<td>44 0.0501</td>
<td>5 0.0226</td>
<td>15 0.0752</td>
<td>25 0.2508</td>
<td>35 0.0501</td>
<td>45 0.0527</td>
<td>6 0.1504</td>
<td>16 0.3009</td>
</tr>
<tr>
<td>7 0.0501</td>
<td>17 0.0752</td>
<td>27 0.0752</td>
<td>37 0.0752</td>
<td>47 0.1253</td>
<td>8 0.0250</td>
<td>18 0.0376</td>
<td>28 0.1002</td>
<td>38 0.0501</td>
<td>48 0.1053</td>
<td>9 0.0150</td>
<td>19 0.0501</td>
</tr>
<tr>
<td>10 0.2005</td>
<td>20 0.0226</td>
<td>30 0.0978</td>
<td>40 0.0602</td>
<td>50 0.0301</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*A sample of customers served at the copying machine was taken for five consecutive 9-hour work days plus 5 hours on Saturday for a total of 50 observations. The average service time is 0.0917 hour per copying job, or 5.499 minutes per job. The equivalent service rate is 10.91 jobs per hour (that is, 10.91 jobs/hour = (60 minutes/hour)/5.5 minutes/job).
past school projects had suggested using queuing models for one project and simulation for the other project. Gifford was never clear on how the method of analysis was chosen. Now, she wondered which methodology she should use for the Bourbon County Court situation.

To organize her thinking Gifford listed a few of the questions she needed to address as follows:

1. Assuming a Poisson arrival distribution and an exponential service time distribution, apply queuing models to the case situation and evaluate the results.

2. What are the economics of the situation using queuing model analysis?

3. What are your final recommendations using queuing model analysis?

4. Advanced Assignment (requires the use of Crystal Ball on the CD-ROM). Do the customer arrival and service empirical (actual) distributions in the case match the theoretical distributions assumed in queuing models?

---

ENDNOTES

1 Brady, D., “Why Service Stinks,” Business Week October 23, 2000, pp. 118–128. This episode is partially based on this article.


