# Learning Objectives

After completing this chapter, students will be able to:

1. Describe the trade-off curves for cost of waiting time and cost of service.
2. Understand the three parts of a queuing system: the calling population, the queue itself, and the service facility.
3. Describe the basic queuing system configurations.
4. Understand the assumptions of the common models dealt with in this chapter.
5. Analyze a variety of operating characteristics of waiting lines.
Introduction

- Queuing theory is the study of waiting lines.
- It is one of the oldest and most widely used quantitative analysis techniques.
- The three basic components of a queuing process are arrivals, service facilities, and the actual waiting line.
- Analytical models of waiting lines can help managers evaluate the cost and effectiveness of service systems.

Waiting Line Costs

- Most waiting line problems are focused on finding the ideal level of service a firm should provide.
- In most cases, this service level is something management can control.
- When an organization does have control, they often try to find the balance between two extremes.
Waiting Line Costs

- There is generally a trade-off between cost of providing service and cost of waiting time.
  - A large staff and many service facilities generally result in high levels of service but have high costs.
  - Having the minimum number of service facilities keeps service cost down but may result in dissatisfied customers.
- Service facilities are evaluated on their total expected cost, which is the sum of service costs and waiting costs.
- Organizations typically want to find the service level that minimizes the total expected cost.

Queuing Costs and Service Levels

![Figure 13.1](https://via.placeholder.com/150)

**Figure 13.1**

Cost of Providing Service

Cost of Waiting Time

Optimal Service Level

Total Expected Cost

Service Level
Three Rivers Shipping Company

Three Rivers Shipping operates a docking facility on the Ohio River.

An average of 5 ships arrive to unload their cargos each shift.

Idle ships are expensive.

More staff can be hired to unload the ships, but that is expensive as well.

Three Rivers Shipping Company wants to determine the optimal number of teams of stevedores to employ each shift to obtain the minimum total expected cost.

### Waiting Line Cost Analysis

<table>
<thead>
<tr>
<th>NUMBER OF TEAMS OF STEVEDORES WORKING</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Average number of ships arriving per-shift</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>(b) Average time each ship waits to be unloaded (hours)</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>(c) Total ship hours lost per shift (a x b)</td>
<td>35</td>
<td>20</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>(d) Estimated cost per hour of idle ship time</td>
<td>$1,000</td>
<td>$1,000</td>
<td>$1,000</td>
<td>$1,000</td>
</tr>
<tr>
<td>(e) Value of ship’s lost time or waiting cost (c x d)</td>
<td>$35,000</td>
<td>$20,000</td>
<td>$15,000</td>
<td>$10,000</td>
</tr>
<tr>
<td>(f) Estimated cost per ship: stevedore team salary or service cost</td>
<td>$6,000</td>
<td>$12,000</td>
<td>$18,000</td>
<td>$24,000</td>
</tr>
<tr>
<td>(g) Total expected cost (e + f)</td>
<td>$41,000</td>
<td>$32,000</td>
<td>$33,000</td>
<td>$34,000</td>
</tr>
</tbody>
</table>

Table 13.1
Characteristics of a Queuing System

There are three parts to a queuing system:
1. The arrivals or inputs to the system (sometimes referred to as the calling population).
2. The queue or waiting line itself.
3. The service facility.

These components have their own characteristics that must be examined before mathematical models can be developed.

Characteristics of a Queuing System

Arrival Characteristics have three major characteristics: size, pattern, and behavior.
- The size of the calling population can be either unlimited (essentially infinite) or limited (finite).
- The pattern of arrivals can arrive according to a known pattern or can arrive randomly.
- Random arrivals generally follow a Poisson distribution.
Characteristics of a Queuing System

Behavior of arrivals

- Most queuing models assume customers are patient and will wait in the queue until they are served and do not switch lines.
- Balking refers to customers who refuse to join the queue.
- Reneging customers enter the queue but become impatient and leave without receiving their service.
- That these behaviors exist is a strong argument for the use of queuing theory to managing waiting lines.

Characteristics of a Queuing System

Waiting Line Characteristics

- Waiting lines can be either limited or unlimited.
- Queue discipline refers to the rule by which customers in the line receive service.
  - The most common rule is first-in, first-out (FIFO).
  - Other rules are possible and may be based on other important characteristics.
- Other rules can be applied to select which customers enter which queue, but may apply FIFO once they are in the queue.
Characteristics of a Queuing System

Service Facility Characteristics

- Basic queuing system configurations:
  - Service systems are classified in terms of the number of channels, or servers, and the number of phases, or service stops.
  - A single-channel system with one server is quite common.
  - Multichannel systems exist when multiple servers are fed by one common waiting line.
  - In a single-phase system, the customer receives service from just one server.
  - In a multiphase system, the customer has to go through more than one server.

---

Four basic queuing system configurations

- Single-Channel, Single-Phase System
- Single-Channel, Multiphase System
- Multichannel, Single-Phase System
- Multichannel, Multiphase System

Figure 13.2
Characteristics of a Queuing System

Service time distribution
- Service patterns can be either constant or random.
- Constant service times are often machine controlled.
- More often, service times are randomly distributed according to a negative exponential probability distribution.
- Analysts should observe, collect, and plot service time data to ensure that the observations fit the assumed distributions when applying these models.

Identifying Models Using Kendall Notation

- D. G. Kendall developed a notation for queuing models that specifies the pattern of arrival, the service time distribution, and the number of channels.
- Notation takes the form:
  \[ \text{Arrival distribution} / \text{Service time distribution} / \text{Number of service channels open} \]
- Specific letters are used to represent probability distributions:
  - \( M \) = Poisson distribution for number of occurrences
  - \( D \) = constant (deterministic) rate
  - \( G \) = general distribution with known mean and variance
Identifying Models Using Kendall Notation

- A single-channel model with Poisson arrivals and exponential service times would be represented by:
  \[ M/M/1 \]
- If a second channel is added the notation would read:
  \[ M/M/2 \]
- A three-channel system with Poisson arrivals and constant service time would be
  \[ M/M/3 \]
- A four-channel system with Poisson arrivals and normally distributed service times would be
  \[ M/M/4 \]

Single-Channel Model, Poisson Arrivals, Exponential Service Times (M/M/1)

Assumptions of the model:
- Arrivals are served on a FIFO basis.
- There is no balking or reneging.
- Arrivals are independent of each other but the arrival rate is constant over time.
- Arrivals follow a Poisson distribution.
- Service times are variable and independent but the average is known.
- Service times follow a negative exponential distribution.
- Average service rate is greater than the average arrival rate.
Single-Channel Model, Poisson Arrivals, Exponential Service Times (M/M/1)

- When these assumptions are met, we can develop a series of equations that define the queue's operating characteristics.

Queueing Equations:

Let

\[ \lambda = \text{mean number of arrivals per time period} \]
\[ \mu = \text{mean number of customers or units served per time period} \]

The arrival rate and the service rate must be defined for the same time period.

1. The average number of customers or units in the system, \( L \):

\[ L = \frac{\lambda}{\mu - \lambda} \]

2. The average time a customer spends in the system, \( W \):

\[ W = \frac{1}{\mu - \lambda} \]

3. The average number of customers in the queue, \( L_q \):

\[ L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} \]
Single-Channel Model, Poisson Arrivals, Exponential Service Times (M/M/1)

4. The average time a customer spends waiting in the queue, $W_q$:

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

5. The utilization factor for the system, $\rho$, the probability the service facility is being used:

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

6. The percent idle time, $P_0$, or the probability no one is in the system:

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

7. The probability that the number of customers in the system is greater than $k$, $P_{n>k}$:

$$P_{n>k} = \frac{1}{\mu} \left( \frac{\lambda}{\mu} \right)^{k+1}$$
Arnold’s Muffler Shop

Arnold’s mechanic can install mufflers at a rate of 3 per hour.

Customers arrive at a rate of 2 per hour.

So:

\( \lambda = 2 \) cars arriving per hour
\( \mu = 3 \) cars serviced per hour

\[ L = \frac{\lambda}{\mu - \lambda} = \frac{2}{3 - 2} = 1 \] 2 cars in the system on average

\[ W = \frac{1}{\mu - \lambda} = \frac{1}{3 - 2} = 1 \] 1 hour that an average car spends in the system

Arnold’s Muffler Shop

\[ L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{2^2}{3(3 - 2)} = \frac{4}{3} = 1.33 \] cars waiting in line on average

\[ W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{2}{3} \] hour = 40 minutes average waiting time per car

\[ \rho = \frac{\lambda}{\mu} = \frac{2}{3} = 0.67 \] = percentage of time mechanic is busy

\[ P_0 = 1 - \frac{\lambda}{\mu} = 1 - \frac{2}{3} = 0.33 \] = probability that there are 0 cars in the system
Arnold’s Muffler Shop

Probability of more than \( k \) cars in the system

\[
P_{\text{more}} = \left( \frac{2}{3} \right)^{k+1}
\]

Note that this is equal to \( 1 - P_0 = 1 - 0.33 = 0.667 \)

Note that this is equal to \( 1 - 0.667 = 0.33 \)

Implies that there is a 19.3% chance that more than 3 cars are in the system

\begin{align*}
&k & P_{\text{more}} \\
&0 & 0.667 \\
&1 & 0.444 \\
&2 & 0.296 \\
&3 & 0.198 \\
&4 & 0.132 \\
&5 & 0.088 \\
&6 & 0.058 \\
&7 & 0.039 \\
\end{align*}

Excel QM Solution to Arnold’s Muffler Example

Program 13.1
Arnold’s Muffler Shop

Introducing costs into the model:

- Arnold wants to do an economic analysis of the queuing system and determine the waiting cost and service cost.
- The total service cost is:
  \[
  \text{Total service cost} = m \cdot c_s
  \]

Waiting cost when the cost is based on time in the system:

- Total waiting cost = (total time spent waiting by all arrivals) \times (Cost of waiting) = \( \lambda \cdot W \cdot c_w \)
- Total waiting cost = (Number of arrivals) \times (Average wait per arrival) \cdot c_w

If waiting time cost is based on time in the queue:

- Total waiting cost = \( \lambda \cdot W_q \cdot c_w \)
Arnold's Muffler Shop

So the total cost of the queuing system when based on time in the system is:

Total cost = Total service cost + Total waiting cost

Total cost = \( mC_s + \lambda W C_w \)

And when based on time in the queue:

Total cost = \( mC_s + \lambda W_q C_w \)

Arnold estimates the cost of customer waiting time in line is $50 per hour.

Total daily waiting cost = (8 hours per day) \( \lambda W_q C_w \)
= (8)(2)(\frac{2}{3})(50) = $533.33

Arnold has identified the mechanics wage $7 per hour as the service cost.

Total daily service cost = (8 hours per day) \( mC_s \)
= (8)(15) = $120

So the total cost of the system is:

Total daily cost of the queuing system = $533.33 + $120 = $653.33
Arnold’s Muffler Shop

- Arnold is thinking about hiring a different mechanic who can install mufflers at a faster rate.
- The new operating characteristics would be:
  \[
  \lambda = 2 \text{ cars arriving per hour} \\
  \mu = 4 \text{ cars serviced per hour}
  \]

\[
L = \frac{\lambda}{\mu + \lambda} = \frac{2}{4 - 2} = 1 \text{ car in the system on the average}
\]

\[
W = \frac{1}{\mu - \lambda} = \frac{1}{4 - 2} = 1/2 \text{ hour that an average car spends in the system}
\]

Arnold’s Muffler Shop

\[
L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{2^2}{8(4)} = 0.5 = 1/2 \text{ car waiting in line on the average}
\]

\[
W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{1}{4} \text{ hour} = 15 \text{ minutes average waiting time per car}
\]

\[
\rho = \frac{\lambda}{\mu} = \frac{2}{4} = 0.5 = \text{ percentage of time mechanic is busy}
\]

\[
P_0 = 1 - \frac{\lambda}{\mu} = 1 - \frac{2}{4} = 0.5 = \text{ probability that there are 0 cars in the system}
\]
Arnold’s Muffler Shop

Probability of more than \( k \) cars in the system

\[
P_{out} = \left( \frac{1}{\lambda} \right)^{k+1}
\]

<table>
<thead>
<tr>
<th>( k )</th>
<th>( P_{out} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.300</td>
</tr>
<tr>
<td>1</td>
<td>0.250</td>
</tr>
<tr>
<td>2</td>
<td>0.125</td>
</tr>
<tr>
<td>3</td>
<td>0.062</td>
</tr>
<tr>
<td>4</td>
<td>0.031</td>
</tr>
<tr>
<td>5</td>
<td>0.016</td>
</tr>
<tr>
<td>6</td>
<td>0.008</td>
</tr>
<tr>
<td>7</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Arnold’s Muffler Shop Case

- The customer waiting cost is the same $50 per hour:
  \[
  \text{Total daily waiting cost} = (8 \text{ hours per day}) \lambda W C_w
  = (8)(2)(1/4)($50) = $200.00
  \]

- The new mechanic is more expensive at $20 per hour:
  \[
  \text{Total daily service cost} = (8 \text{ hours per day}) m \bar{C}_s
  = (8)(1)($20) = $160
  \]

- So the total cost of the system is:
  \[
  \text{Total daily cost of the queuing system} = $200 + $160 = $360
  \]
Arnold’s Muffler Shop

- The total time spent waiting for the 16 customers per day was formerly:
  \[(16 \text{ cars per day}) \times (\frac{2}{3} \text{ hour per car}) = 10.67 \text{ hours}\]
- It is now:
  \[(16 \text{ cars per day}) \times (\frac{1}{4} \text{ hour per car}) = 4 \text{ hours}\]
- The total daily system costs are less with the new mechanic resulting in significant savings:
  \[\$653.33 - \$360 = \$293.33\]

Enhancing the Queuing Environment

- Reducing waiting time is not the only way to reduce waiting cost.
- Reducing the unit waiting cost \(c_w\) will also reduce total waiting cost.
- This might be less expensive to achieve than reducing either \(W\) or \(W_q\).
Assumptions of the model:
- Arrivals are served on a FIFO basis.
- There is no balking or reneging.
- Arrivals are independent of each other but the arrival rate is constant over time.
- Arrivals follow a Poisson distribution.
- Service times are variable and independent but the average is known.
- Service times follow a negative exponential distribution.
- The average service rate is greater than the average arrival rate.

Equations for the multichannel queuing model:
- Let
  - $m$ = number of channels open
  - $\lambda$ = average arrival rate
  - $\mu$ = average service rate at each channel

1. The probability that there are zero customers in the system is:

$$ P_0 = \sum_{n=0}^{m-1} \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!} + \frac{1}{m \left(\frac{\lambda}{\mu}\right) m \mu - \lambda} \quad \text{for } m \mu > \lambda $$
### Multichannel Model, Poisson Arrivals, Exponential Service Times (M/M/m)

2. The average number of customers or units in the system

\[
L = \frac{\lambda \mu (\lambda/\mu)^m}{(m-1)! (m \mu - \lambda)^2} P_0 + \frac{1}{\mu} \frac{\lambda}{\mu} 
\]

3. The average time a unit spends in the waiting line or being served, in the system

\[
W = \frac{\mu (\lambda/\mu)^m}{(m-1)! (m \mu - \lambda)^2} P_0 + \frac{1}{\mu} \frac{1}{\lambda} \frac{L}{\mu} 
\]

### Multichannel Model, Poisson Arrivals, Exponential Service Times (M/M/m)

4. The average number of customers or units in line waiting for service

\[
L_q = L - \frac{\lambda}{\mu} 
\]

5. The average number of customers or units in line waiting for service

\[
W_q = W - \frac{1}{\mu} = \frac{L_q}{\lambda} 
\]

6. The average number of customers or units in line waiting for service

\[
\rho = \frac{\lambda}{m \mu} 
\]
Arnold’s Muffler Shop Revisited

- Arnold wants to investigate opening a second garage bay.
- He would hire a second worker who works at the same rate as his first worker.
- The customer arrival rate remains the same.

\[ P = \frac{1}{m!} \frac{\lambda^m}{\mu^m} \]  
for \( m \mu > \lambda 

\[ P_0 = \frac{1}{\sum_{n=0}^{\infty} \frac{1}{n!} \frac{\lambda^n}{\mu^n} \frac{\mu}{m \mu - \lambda}} = 0.5 \]

= Probability of 0 cars in the system

Arnold’s Muffler Shop Revisited

- Average number of cars in the system

\[ L = \frac{\lambda \mu (\lambda / \mu)^m}{(m - 1) (m \mu - \lambda)^2} P_0 + \frac{\lambda}{\mu} \]

\[ L = \frac{2(3)! (2) (3)!}{(1)! (2)(3) - 2!} \left( \frac{1}{2} \right) + \frac{2}{3} = 0.75 \]

- Average time a car spends in the system

\[ W = \frac{L}{\lambda} = \frac{3}{8}\text{ hour} = 22\frac{1}{2}\text{ minutes} \]
Arnold’s Muffler Shop Revisited

- Average number of cars in the queue
  \[ L_q = L - \frac{\lambda}{\mu} = \frac{3}{4} - \frac{2}{3} = \frac{1}{12} = 0.083 \]

- Average time a car spends in the queue
  \[ W_q = W - \frac{1}{\mu} = \frac{L_q}{\lambda} = \frac{0.083}{2} \approx 0.0415 \text{ hours} = 2 \frac{1}{2} \text{ minutes} \]

Arnold’s Muffler Shop Revisited

- Adding the second service bay reduces the waiting time in line but will increase the service cost as a second mechanic needs to be hired.

  \[ \text{Total daily waiting cost} = (8 \text{ hours per day}) \lambda W_q \]
  \[ = (8)(2)(0.0415)(50) = 33.20 \]

  \[ \text{Total daily service cost} = (8 \text{ hours per day}) \mu c_s \]
  \[ = (8)(2)(15) = 240 \]

- So the total cost of the system is

  \[ \text{Total system cost} = 33.20 + 240 = 273.20 \]

- This is the cheapest option: open the second bay and hire a second worker at the same $15 rate.
## Effect of Service Level on Arnold's Operating Characteristics

<table>
<thead>
<tr>
<th>OPERATING CHARACTERISTIC</th>
<th>LEVEL OF SERVICE</th>
<th>ONE MECHANIC $\mu = 3$</th>
<th>TWO MECHANICS $\mu = 3$ FOR BOTH</th>
<th>ONE FAST MECHANIC $\mu = 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability that the system is empty $P_0$</td>
<td>0.33</td>
<td>0.50</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Average number of cars in the system $(L)$</td>
<td>2 cars</td>
<td>0.75 cars</td>
<td>1 car</td>
<td></td>
</tr>
<tr>
<td>Average time spent in the system $(W)$</td>
<td>60 minutes</td>
<td>22.5 minutes</td>
<td>30 minutes</td>
<td></td>
</tr>
<tr>
<td>Average number of cars in the queue $(L_q)$</td>
<td>1.33 cars</td>
<td>0.083 car</td>
<td>0.30 car</td>
<td></td>
</tr>
<tr>
<td>Average time spent in the queue $(W_q)$</td>
<td>80 minutes</td>
<td>2.5 minutes</td>
<td>15 minutes</td>
<td></td>
</tr>
</tbody>
</table>

Table 13.2

## Excel QM Solution to Arnold's Muffler Multichannel Example

Program 13.2

<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>Arnold's Muffler Shop Multichannel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Waiting Lines</td>
<td>M/M/1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>The arrival RATE and service RATE both must be rates and use the same time unit. Given is</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Data</td>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Arrival rate $(\lambda)$</td>
<td>2</td>
<td>Average server utilization $(\rho)$</td>
<td>0.333333</td>
</tr>
<tr>
<td>8</td>
<td>Service rate $(\mu)$</td>
<td>3</td>
<td>Average number of customers in the queue $(L_q)$</td>
<td>0.083333</td>
</tr>
<tr>
<td>9</td>
<td>Number of servers $(s)$</td>
<td>2</td>
<td>Average number of customers in the system $(L)$</td>
<td>0.75</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td>Average waiting time in the queue $(W_q)$</td>
<td>0.04167</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td>Average time in the system $(W)$</td>
<td>0.375</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td>Probability (% of time) system is empty $(P_0)$</td>
<td>0.5</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td>Probability</td>
<td>Cumulative Probability</td>
</tr>
<tr>
<td>14</td>
<td>Number</td>
<td>Probability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0.500000</td>
<td>0.500000</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>0.333333</td>
<td>0.833333</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>0.111111</td>
<td>0.944444</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>3</td>
<td>0.033333</td>
<td>0.966667</td>
<td></td>
</tr>
</tbody>
</table>
Constant Service Time Model (M/D/1)

- Constant service times are used when customers or units are processed according to a fixed cycle.
- The values for $L_q$, $W_q$, $L$, and $W$ are always less than they would be for models with variable service time.
- In fact both average queue length and average waiting time are halved in constant service rate models.

1. Average length of the queue

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

2. Average waiting time in the queue

$$W_q = \frac{\lambda}{2\mu(\mu - \lambda)}$$
Constant Service Time Model (M/D/1)

3. Average number of customers in the system

\[ L = l_q + \frac{1}{\mu} \]

4. Average time in the system

\[ W = W_q + \frac{1}{\mu} \]

Garcia-Golding Recycling, Inc.

- The company collects and compacts aluminum cans and glass bottles.
- Trucks arrive at an average rate of 8 per hour (Poisson distribution).
- Truck drivers wait about 15 minutes before they empty their load.
- Drivers and trucks cost $60 per hour.
- A new automated machine can process truckloads at a constant rate of 12 per hour.
- A new compactor would be amortized at $3 per truck unloaded.
## Constant Service Time Model (M/D/1)

**Analysis of cost versus benefit of the purchase**

- **Current waiting cost/trip**
  
  \[
  \text{Current waiting cost/trip} = \left( \frac{1}{4} \text{ hour waiting time} \right) \left( \$60/\text{hour cost} \right) = 15 \text{$/trip}
  \]

- **New system**
  
  \[
  \lambda = 8 \text{ trucks/hour arriving} \quad \mu = 12 \text{ trucks/hour served}
  \]

  \[
  W_q = \frac{A}{2\mu(\mu - \lambda)} - \frac{A}{2(12)(12 - 8)} = \frac{1}{12} \text{ hour}
  \]

  \[
  \text{Waiting cost/trip with new compactor} = \left( \frac{1}{12} \text{ hour wait} \right) \left( \$60/\text{hour cost} \right) = 5 \text{$/trip}
  \]

  **Savings**

  \[
  \text{Savings with new equipment} = 15 \text{(current system)} - 5 \text{(new system)} = 10 \text{$/trip}
  \]

  **Cost of new equipment amortized**

  \[
  \text{amortized} = 3 \text{$/trip}
  \]

  **Net savings**

  \[
  \text{Net savings} = 7 \text{$/trip}
  \]

---

### Excel QM Solution for Constant Service Time Model with Garcia-Golding Recycling Example

<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>Garcia-Golding Recycling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Waiting Lines</td>
<td>M/D/1 (Constant Service Times)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>The arrival RATE and service RATE both must be rates and use the same time unit. Given a time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Data</td>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Arrival rate (λ)</td>
<td>8</td>
<td>Average server utilization(ρ)</td>
<td>0.66667</td>
</tr>
<tr>
<td>7</td>
<td>Service rate (μ)</td>
<td>12</td>
<td>Average number of customers in the queue(Lq)</td>
<td>0.66667</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td>Average number of customers in the system(L)</td>
<td>1.3333</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td>Average waiting time in the queue(Wq)</td>
<td>0.08333</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td>Average time in the system(W)</td>
<td>0.16667</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td>Probability (% of time) system is empty (P0)</td>
<td>0.3333</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Program 13.3
Finite Population Model
(M/M/1 with Finite Source)

- When the population of potential customers is limited, the models are different.
- There is now a dependent relationship between the length of the queue and the arrival rate.
- The model has the following assumptions:
  1. There is only one server.
  2. The population of units seeking service is finite.
  3. Arrivals follow a Poisson distribution and service times are exponentially distributed.
  4. Customers are served on a first-come, first-served basis.

Examples

- Equipment repair in a company has five machines.
- Maintenance for a fleet of 10 commuter airplanes.
- Hospital ward has 20 beds
Finite Population Model
(M/M/1 with Finite Source)

Equations for the finite population model:

Using $\lambda$ = mean arrival rate, $\mu$ = mean service rate, and $N$ = size of the population, the operating characteristics are:

1. Probability that the system is empty:

   $$ P_0 = \frac{1}{\sum_{n=0}^{N-1} \frac{N!}{(N-n)!} \left(\frac{\lambda}{\mu}\right)^n} $$

2. Average length of the queue:

   $$ L_q = N + \left(\frac{\lambda}{\mu}\right) \left(1 - P_0\right) $$

3. Average number of customers (units) in the system:

   $$ L = L_q + (1 - P_0) $$

4. Average waiting time in the queue:

   $$ W_q = \frac{L_q}{(N - L)\lambda} $$
Finite Population Model
(M/M/1 with Finite Source)

5. Average time in the system:

\[ W = W_q + \frac{1}{\mu} \]

6. Probability of \( n \) units in the system:

\[ p_n = \frac{N!}{(N-n)!} \left( \frac{\lambda}{\mu} \right)^n p_0 \text{ for } n = 0, 1, ..., N \]

Department of Commerce

- The Department of Commerce has five printers that each need repair after about 20 hours of work.
- Breakdowns will follow a Poisson distribution.
- The technician can service a printer in an average of about 2 hours, following an exponential distribution.
- Therefore:

\[ \lambda = \frac{1}{20} = 0.05 \text{ printer/hour} \]
\[ \mu = \frac{1}{2} = 0.50 \text{ printer/hour} \]
Department of Commerce Example

1. \[ p_0 = \frac{2}{\sum_{i=0}^{5} \frac{5!}{i!(5-i)!}(0.05)^i(0.5)^{5-i}} = 0.564 \]

2. \[ L_0 = 5 + \left( \frac{0.05 + 0.5}{0.05} \right)(1 - p_0) = 0.2 \text{ printer} \]

3. \[ L = 0.2 + (1 - 0.564) = 0.64 \text{ printer} \]

Department of Commerce Example

4. \[ W_f = \frac{0.2}{(5 - 0.64)(0.05)} = \frac{0.2}{0.22} = 0.91 \text{ hour} \]

5. \[ W = 0.9 + \frac{1}{0.5} = 2.91 \text{ hours} \]

If printer downtime costs $120 per hour and the technician is paid $25 per hour, the total cost is:

\[ \text{Total hourly cost} = \frac{(\text{Average number of printers down})}{(\text{Cost per downtime hour})} \times \text{Cost per technician hour} \]

\[ = (0.64)($120) + 25 = $101.80 \]
Excel QM For Finite Population Model with Department of Commerce Example

<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>Department of Commerce</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Waiting Lines</td>
<td>M/M/s with a finite population</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>The arrival rate is for each member of the population. If they go for service every 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Data</td>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Arrival rate (λ) per customer</td>
<td>0.05</td>
<td>Average server utilization (ρ)</td>
<td>0.43605</td>
</tr>
<tr>
<td>8</td>
<td>Service rate (μ)</td>
<td>0.5</td>
<td>Average number of customers in the queue (Lq)</td>
<td>0.20947</td>
</tr>
<tr>
<td>9</td>
<td>Number of servers</td>
<td>1</td>
<td>Average number of customers in the system (Ls)</td>
<td>0.63952</td>
</tr>
<tr>
<td>10</td>
<td>Population size (N)</td>
<td>5</td>
<td>Average waiting time in the queue (Wq)</td>
<td>0.93326</td>
</tr>
<tr>
<td>11</td>
<td>Average time in the system (Ws)</td>
<td>2.93326</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Probability (% of time) system is empty (P0)</td>
<td>0.56395</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Effective arrival rate</td>
<td></td>
<td>0.21502</td>
<td></td>
</tr>
</tbody>
</table>

Program 13.4

Some General Operating Characteristic Relationships

- Certain relationships exist among specific operating characteristics for any queuing system in a steady state.
- A steady state condition exists when a system is in its normal stabilized condition, usually after an initial transient state.
- The first of these are referred to as Little's Flow Equations:
  \[ L = \lambda W \]
  \[ L_s = \lambda W_s \]
  \[ W = W_s + 1/\mu \]
More Complex Queuing Models and the Use of Simulation

- In the real world there are often variations from basic queuing models.
- Computer simulation can be used to solve these more complex problems.
- Simulation allows the analysis of controllable factors.
- Simulation should be used when standard queuing models provide only a poor approximation of the actual service system.