Chapter 12

Waiting Line Analysis for Service Improvement

Operations Management - 5th Edition

Roberta Russell & Bernard W. Taylor, III

Beni Asllani
University of Tennessee at Chattanooga

Lecture Outline

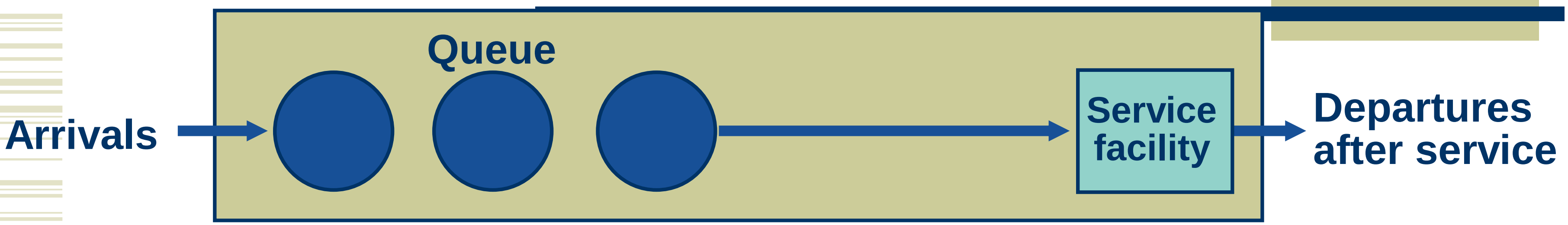
- Elements of Waiting Line Analysis
- Waiting Line Analysis and Quality
- Single Server Models
- Multiple Server Model

Waiting Line Analysis

- Operating characteristics
  - average values for characteristics that describe the performance of a waiting line system
- Queue
  - A single waiting line
- Waiting line system consists of
  - Arrivals
  - Servers
  - Waiting line structures

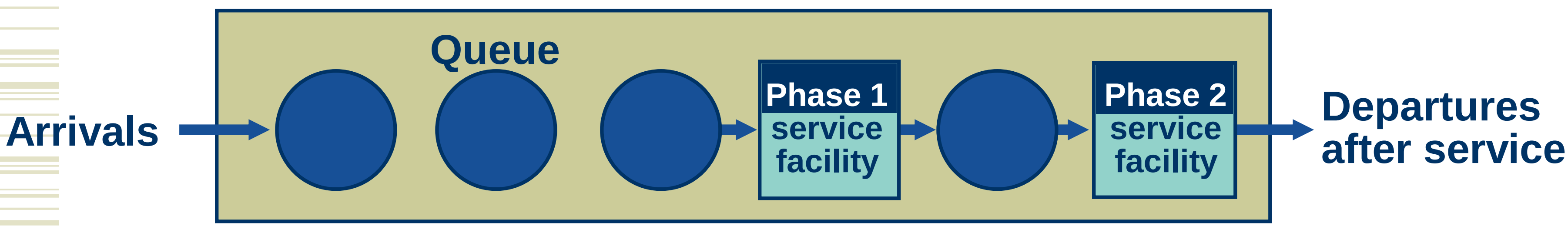
Queuing System Designs

A family dentist's office



Single-channel, single-phase system

A McDonald's dual window drive-through

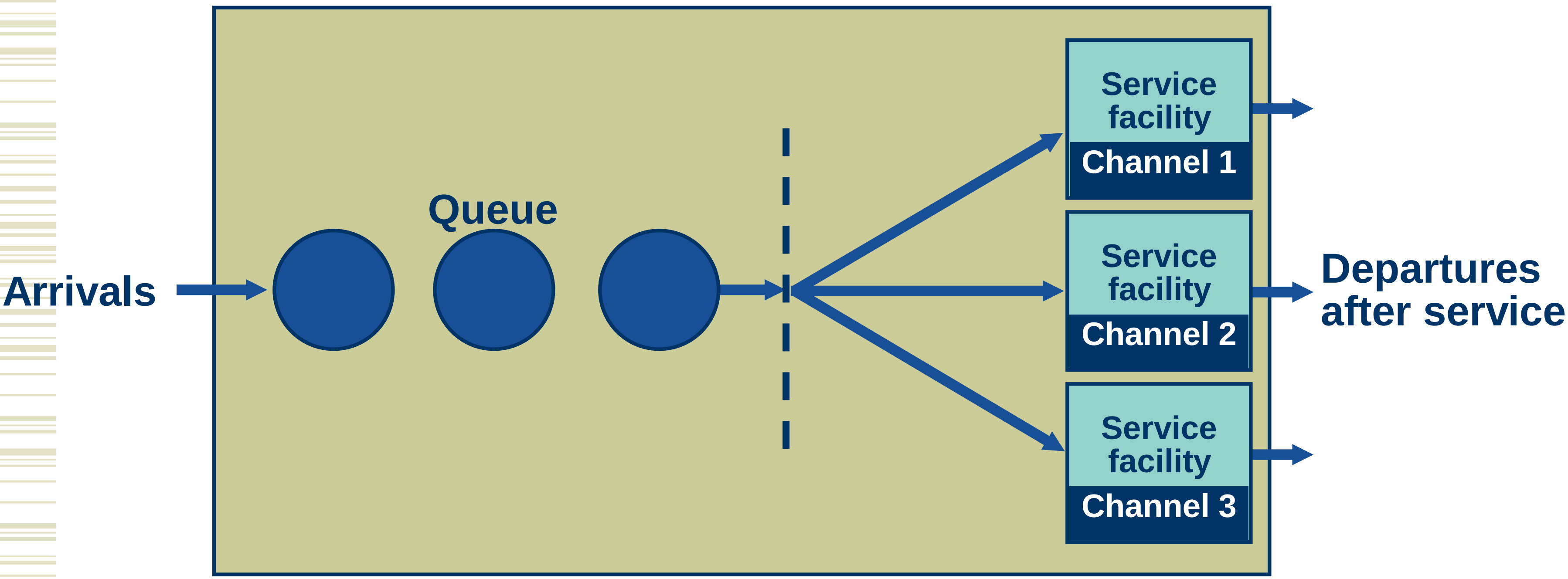


Single-channel, multiphase system

Figure D.3

Queuing System Designs

Most bank and post office service windows

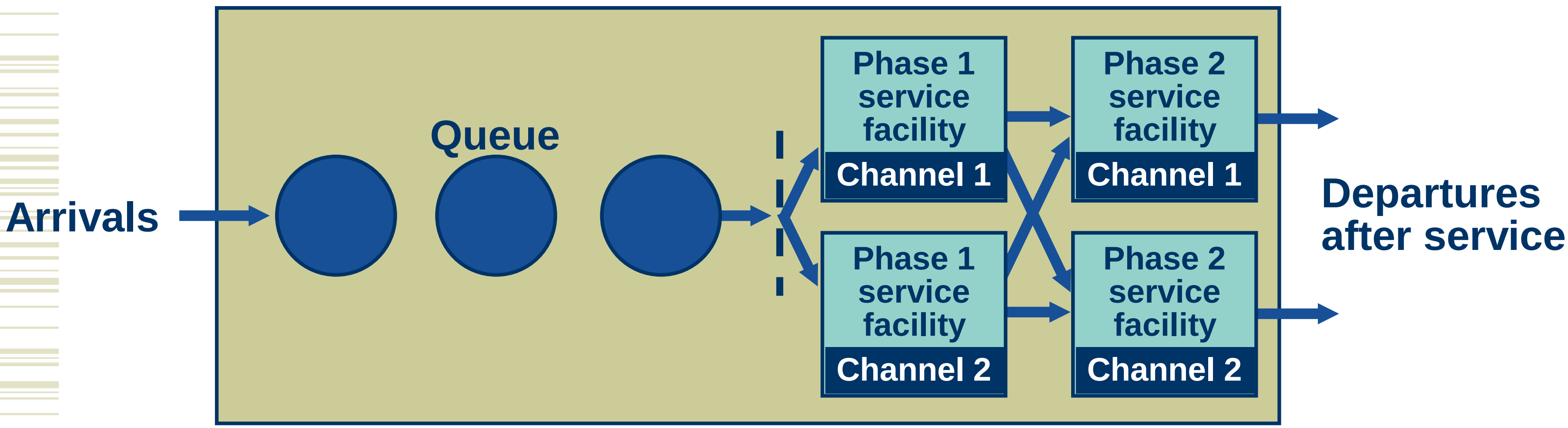


Multi-channel, single-phase system

Figure D.3

Queuing System Designs

Some college registrations



Multi-channel, multiphase system

Figure D.3

Elements of a Waiting Line

- Calling population
  - Source of customers
  - Infinite - large enough that one more customer can always arrive to be served
  - Finite - countable number of potential customers
- Arrival rate (λ)
  - Frequency of customer arrivals at waiting line system
  - Typically follows Poisson distribution

Elements of a Waiting Line (cont.)

- Service time
  - Often follows negative exponential distribution
  - Average service rate = μ
- Arrival rate (λ) must be less than service rate (μ) or system never clears out

Elements of a Waiting Line (cont.)

- Queue discipline
  - Order in which customers are served
  - First come, first served is most common
- Length can be infinite or finite
  - Infinite is most common
  - Finite is limited by some physical

Basic Waiting Line Structures

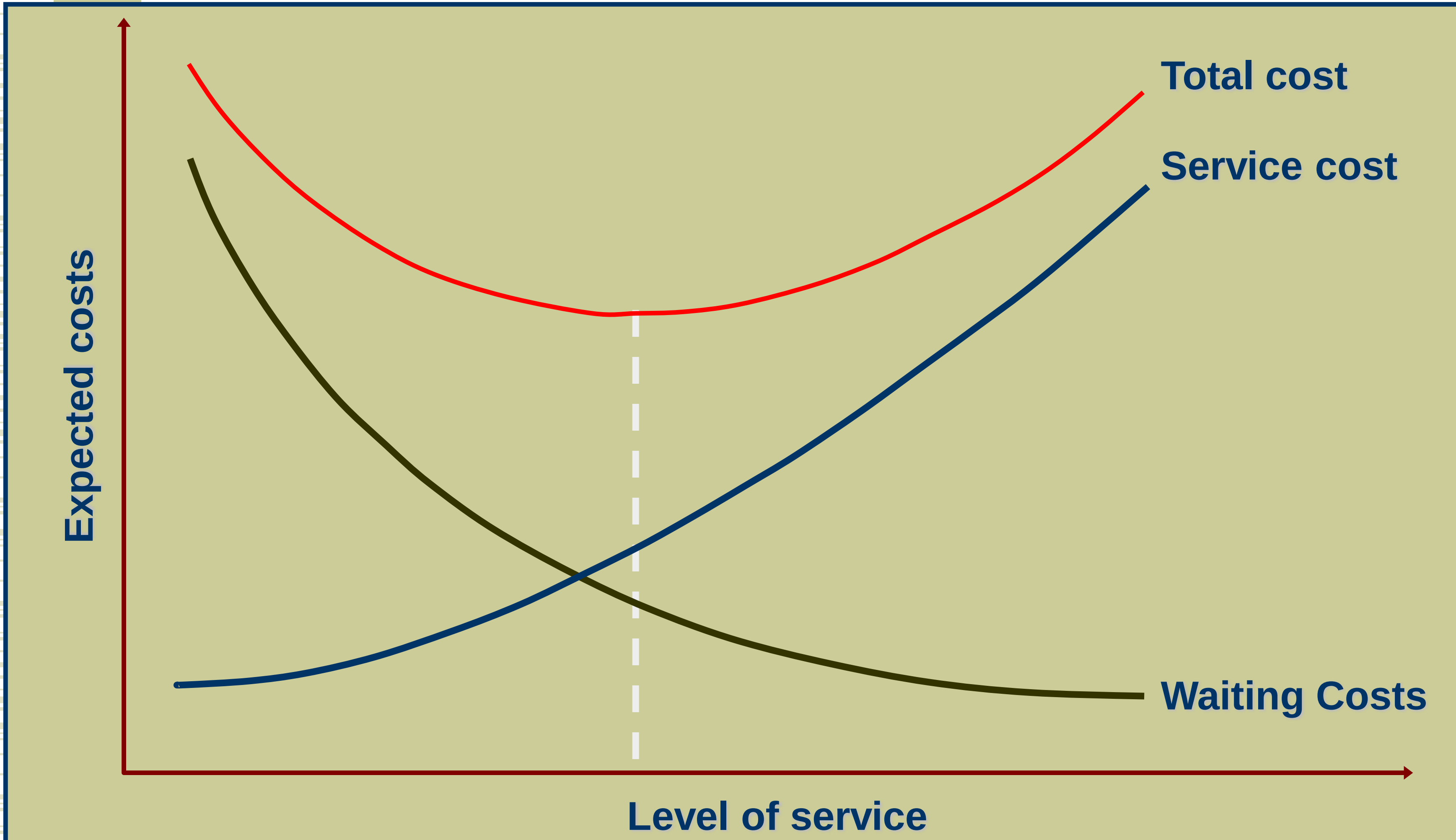
- Channels are the number of parallel servers
  - Single channel
  - Multiple channels
- Phases denote number of sequential servers the customer must go through
  - Single phase
  - Multiple phases
- Steady state
  - A constant, average value for performance characteristics that system will reach after a long time

Operating Characteristics

NOTATION	OPERATING CHARACTERISTIC
L	Average number of customers in the system (waiting and being served)
L_q	Average number of customers in the waiting line
W	Average time a customer spends in the system (waiting and being served)
W_q	Average time a customer spends waiting in line
P_0	Probability of no (zero) customers in the system
P_n	Probability of *n* customers in the system
ρ	Utilization rate; the proportion of time the system is in use

Table 16.1

Cost Relationship in Waiting Line Analysis



Waiting Line Costs and Quality Service

- Traditional view is that the level of service should coincide with minimum point on total cost curve
- TQM approach is that absolute quality service will be the most cost-effective in the long run

Single-server Models

- All assume Poisson arrival rate
- Variations
  - Exponential service times
  - General (or unknown) distribution of service times
  - Constant service times
  - Exponential service times with finite queue length
  - Exponential service times with finite calling population

Basic Single-Server Model: Assumptions

- Poisson arrival rate
- Exponential service times
- First-come, first-served queue discipline
- Infinite queue length
- Infinite calling population
- λ = mean arrival rate
- μ = mean service rate

Formulas for Single-Server Model

Probability that no customers are in the system (either in the queue or being served)

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

Probability of exactly *n* customers in the system

$$P_n = \left(\frac{\lambda}{\mu}\right)^n \cdot P_0$$

$$= \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right)$$

Average number of customers in the system

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

Average number of customers in the waiting line

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

Formulas for Single-Server Model (cont.)

Average time a customer spends in the queuing system

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

Average time a customer spends waiting in line to be served

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

Probability that the server is busy and the customer has to wait

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

Probability that the server is idle and a customer can be served

$$I = 1 - \rho$$

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

A Single-Server Model

Given λ = 24 per hour, μ = 30 customers per hour

Probability of no customers in the system

$$P_0 = \left(1 - \frac{\lambda}{\mu}\right) = \left(1 - \frac{24}{30}\right) = 0.20$$

Average number of customers in the system

$$L = \frac{\lambda}{\mu - \lambda} = \frac{24}{30 - 24} = 4$$

Average number of customers waiting in line

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

A Single-Server Model

Average time in the system per customer

$$W = \frac{1}{\mu - \lambda} = \frac{1}{30 - 24} = 0.167 \text{ hour}$$

Average time waiting in line per customer

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{24}{30(30 - 24)} = 0.133$$

Probability that the server will be busy and the customer must wait

$$\rho = \frac{\lambda}{\mu} = \frac{24}{30} = 0.80$$

Probability the server will be idle

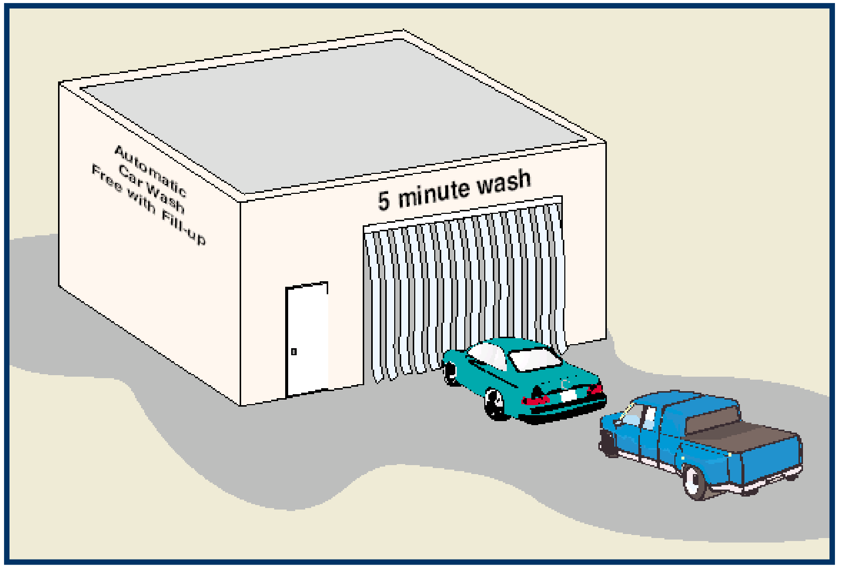
$$I = 1 - \rho = 1 - 0.80 = 0.20$$

Service Improvement Analysis

- Possible Alternatives
  - Another employee to pack up purchases
    - service rate will increase from 30 customers to 40 customers per hour
    - waiting time will reduce to only 2.25 minutes
  - Another checkout counter
    - arrival rate at each register will decrease from 24 to 12 per hour
    - customer waiting time will be 1.33 minutes
- Determining whether these improvements are worth the cost to achieve them is the crux of waiting line analysis

Constant Service Times

- Constant service times occur with machinery and automated equipment
- Constant service times are a special case of the single-server model with *undefined* service times



Operating Characteristics for Constant Service Times

Probability that no customers are in system

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

Average number of customers in queue

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

Average number of customers in system

$$L = L_q + \frac{\lambda}{\mu}$$

Operating Characteristics for Constant Service Times (cont.)

Average time customer spends in queue

$$W_q = \frac{L_q}{\lambda}$$

Average time customer spends in the system

$$W = W_q + \frac{1}{\mu}$$

Probability that the server is busy

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

Constant Service Times: Example

Automated car wash with service time = 4.5 min

Cars arrive at rate λ = 10/hour (Poisson)

μ = 60/4.5 = 13.3/hour

$$L_q = \frac{\lambda^2}{2\mu(\mu - \lambda)} = \frac{(10)^2}{2(13.3)(13.3 - 10)} = 1.14 \text{ cars waiting}$$

$$W_q = \frac{L_q}{\lambda} = 1.14/10 = .114 \text{ hour or } 6.84 \text{ minutes}$$

Finite Queue Length

- A physical limit exists on length of queue
- M = maximum number in queue
- Service rate does not have to exceed arrival rate ($\mu > \lambda$) to obtain steady-state conditions

Probability that no customers are in system

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

Probability of exactly *n* customers in system

$$P_n = (P_0)\left(\frac{\lambda}{\mu}\right)^n \quad \text{for } n \leq M$$

Average number of customers in system

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

Finite Queue Length (cont.)

Let P_M = probability a customer will not join system

Average number of customers in queue

$$L_q = L - \frac{\lambda(1 - P_M)}{\mu}$$

Average time customer spends in system

$$W = \frac{L}{\lambda(1 - P_M)}$$

Average time customer spends in queue

$$W_q = W - \frac{1}{\mu}$$

Finite Queue: Example

First National Bank has waiting space for only 3 drive in window cars. $\lambda = 20$, $\mu = 30$, $M = 4$ cars (1 in service + 3 waiting)

Probability that no cars are in the system

$$P_0 = \frac{1 - \lambda/\mu}{1 - (\lambda/\mu)^{M+1}} = \frac{1 - 20/30}{1 - (20/30)^5} = 0.38$$

Probability of exactly 4 cars in the system

$$P_n = (P_0)\left(\frac{\lambda}{\mu}\right)^{n=M} = (0.38)\left(\frac{20}{30}\right)^4 = 0.076$$

Average number of cars in the system

$$L = \frac{\lambda/\mu}{1 - \lambda/\mu} - \frac{(M+1)(\lambda/\mu)^{M+1}}{1 - (\lambda/\mu)^{M+1}} = 1.24$$

Finite Queue: Example (cont.)

Average number of cars in the queue

$$L_q = L - \frac{\lambda(1 - P_M)}{\mu} = 0.62$$

Average time a car spends in the system

$$W = \frac{L}{\lambda(1 - P_M)} = 0.067 \text{ hr}$$

Average time a car spends in the queue

$$W_q = W - \frac{1}{\mu} = 0.033 \text{ hr}$$

Finite Calling Population

Arrivals originate from a finite (countable) population

N = population size

Probability that no customers are in system

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

Probability of exactly n customers in system

$$P_n = \frac{N!}{(N-n)!}\left(\frac{\lambda}{\mu}\right)^n P_0 \quad \text{where } n = 1, 2, \ldots, N$$

Average number of customers in queue

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

Finite Calling Population (cont.)

Average number of customers in system	$L = L_q + (1 - P_0)$
Average time customer spends in queue	$W_q = \dfrac{L_q}{(N - L)\,\lambda}$
Average time customer spends in system	$W = W_q + \dfrac{1}{\mu}$

Finite Calling Population: Example

20 trucks which operate an average of 200 days before breaking down (λ = 1/200 day = 0.005/day)

Mean repair time = 3.6 days (μ = 1/3.6 day = 0.2778/day)

Probability that no trucks are in the system

$P_0 = 0.652$

Average number of trucks in the queue

$L_q = 0.169$

Average number of trucks in system

$L = 0.169 + (1 - 0.652) = .520$

Average time truck spends in queue

$W_q = 1.74$ days

Average time truck spends in system

$W = 5.33$ days

Basic Multiple-server Model

- Two or more independent servers serve a single waiting line
- Poisson arrivals, exponential service, infinite calling population
- $s\mu > \lambda$

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

- Computing P_0 can be time-consuming.
- Tables can used to find P_0 for selected values of ρ and s.

Basic Multiple-server Model (cont.)

Probability of exactly *n* customers in the system

$$P_n = \begin{cases} \dfrac{1}{s!\, s^{n-s}} \left(\dfrac{\lambda}{\mu}\right)^{n} P_0, & \text{for } n > s \\[2ex] \dfrac{1}{n!} \left(\dfrac{\lambda}{\mu}\right)^{n} P_0, & \text{for } n > s \end{cases}$$

Probability an arriving customer must wait

$$P_w = \frac{1}{s!} \left(\frac{\lambda}{\mu}\right)^{s} \frac{s\mu}{s\mu - \lambda} P_0$$

Average number of customers in system

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

Basic Multiple-server Model (cont.)

Average time customer spends in system

$$W = \frac{L}{\lambda}$$

Average number of customers in queue

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

Average time customer spends in queue

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

Utilization factor

$$\rho = \lambda / s\mu$$

Multiple-Server System: Example

Student Health Service Waiting Room

λ = 10 students per hour

μ = 4 students per hour per service representative

s = 3 representatives

$s\mu = (3)(4) = 12$

Probability no students are in the system $P_0 = 0.045$

Number of students in the service area $L = 6$

Multiple-Server System: Example (cont.)

Waiting time in the service area	$W = L / \lambda = 0.60$
Number of students waiting to be served	$L_q = L - \lambda/\mu = 3.5$
Average time students will wait in line	$W_q = L_q/\lambda = 0.35$ hours
Probability that a student must wait	$P_w = 0.703$

Multiple-Server System: Example (cont.)

- Add a 4th server to improve service
- Recompute operating characteristics
  - P_0 = 0.073 prob of no students
  - L = 3.0 students
  - W = 0.30 hour, 18 min in service
  - L_q = 0.5 students waiting
  - W_q = 0.05 hours, 3 min waiting, versus 21 earlier
  - P_w = 0.31 prob that a student must wait