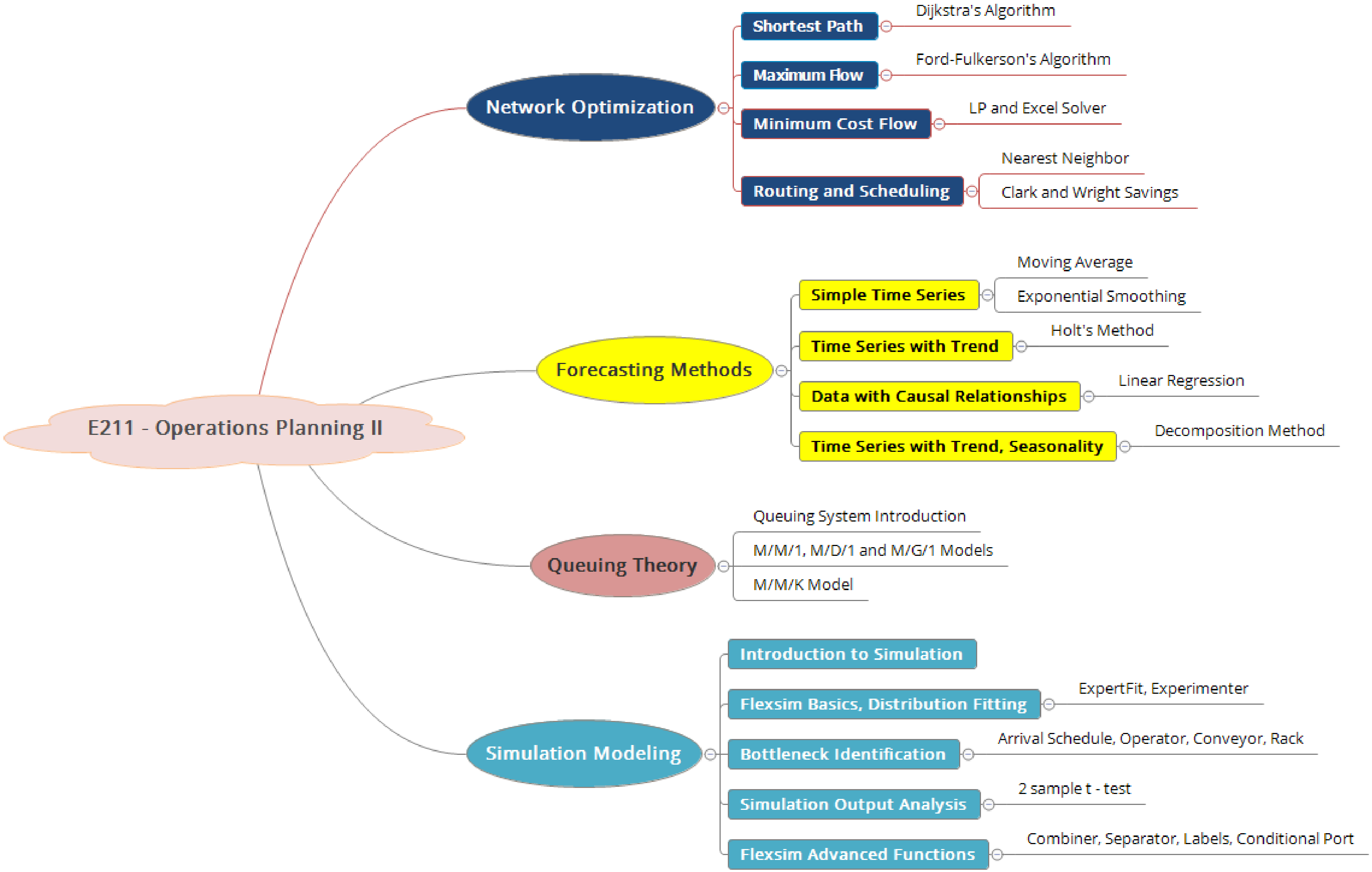
Problem 07
Waiting Line Analysis

E211 – Operations Planning II

SCHOOL OF ENGINEERING

Module Coverage: E211 Topic Tree



Queuing Theory

Queuing Theory – a body of knowledge about waiting lines. It's an important part of operations and a valuable tool for many operations managers.

- It is estimated that the world spent a *billion* hours a year waiting in lines.

- Places we wait in line...

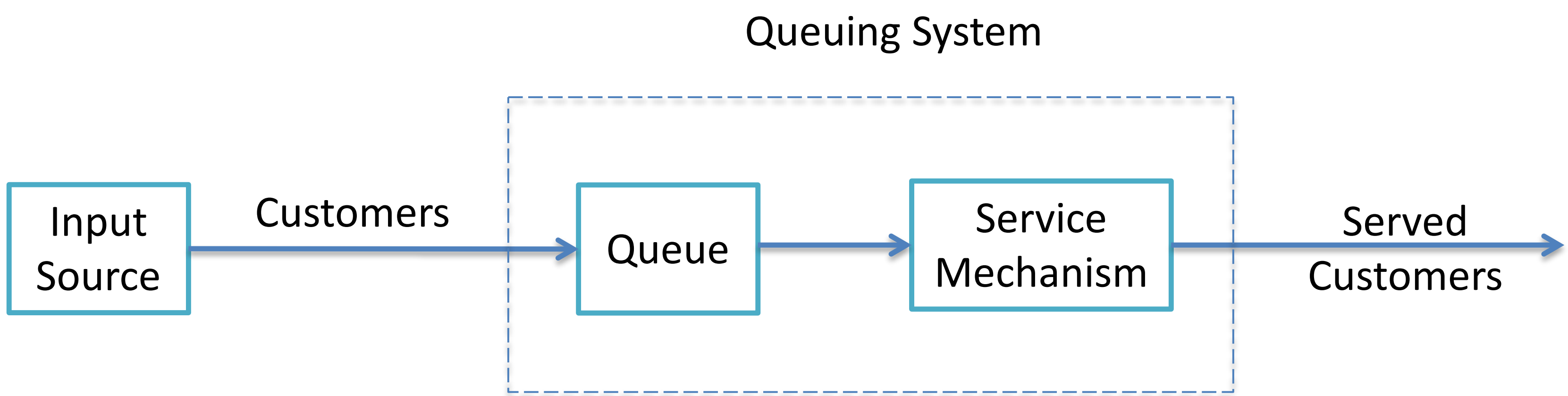
  - Stores
  - Banks
  - Airports
  - Computer websites
  - Hotels
  - Traffic lights
  - Theme parks
  - Post offices
  - Restaurants/Canteen
  - On the phone

Queuing System

A waiting line or queuing system is characterised by **Three** basic components:

- ***Arrival***
  - Customers arrive according to some arrival pattern.
- ***Waiting Line or Queue***
  - Arriving customers may have to wait in one or more queues for service.
- ***Service Mechanism***
  - Server(s)/Counters providing service to customers

The Basic Queuing Process



The basic queuing process:

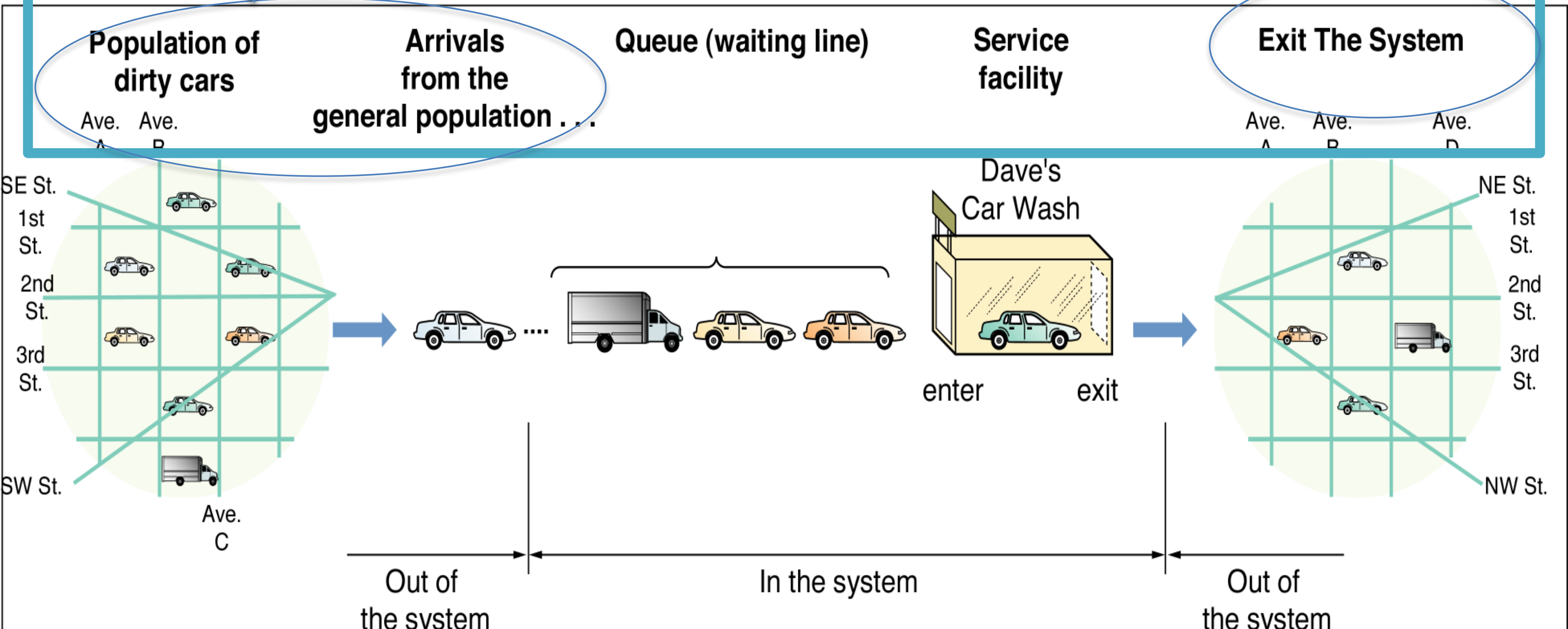
- ***Customers*** requiring service are generated over time by an ***input source***.
- These customers enter the ***queuing system*** and join a ***queue***.
- At certain times, a member of the queue is selected for service by some rule known as the ***queue discipline***.
- The required service is then performed for the customer by the ***service mechanism***, after which the customer leaves the queuing system.

Queuing System Example: Car-Wash

Input to a queuing system

Components of a queuing system

Often not of much interest in queuing theory



Arrival Characteristics

- Size of arrival population
- Behavior of arrivals
- Statistical distribution of arrivals

Waiting-line Characteristics

- Limited vs. Unlimited
- Queue Discipline

Service Characteristics

- Service design
- Statistical distribution of service

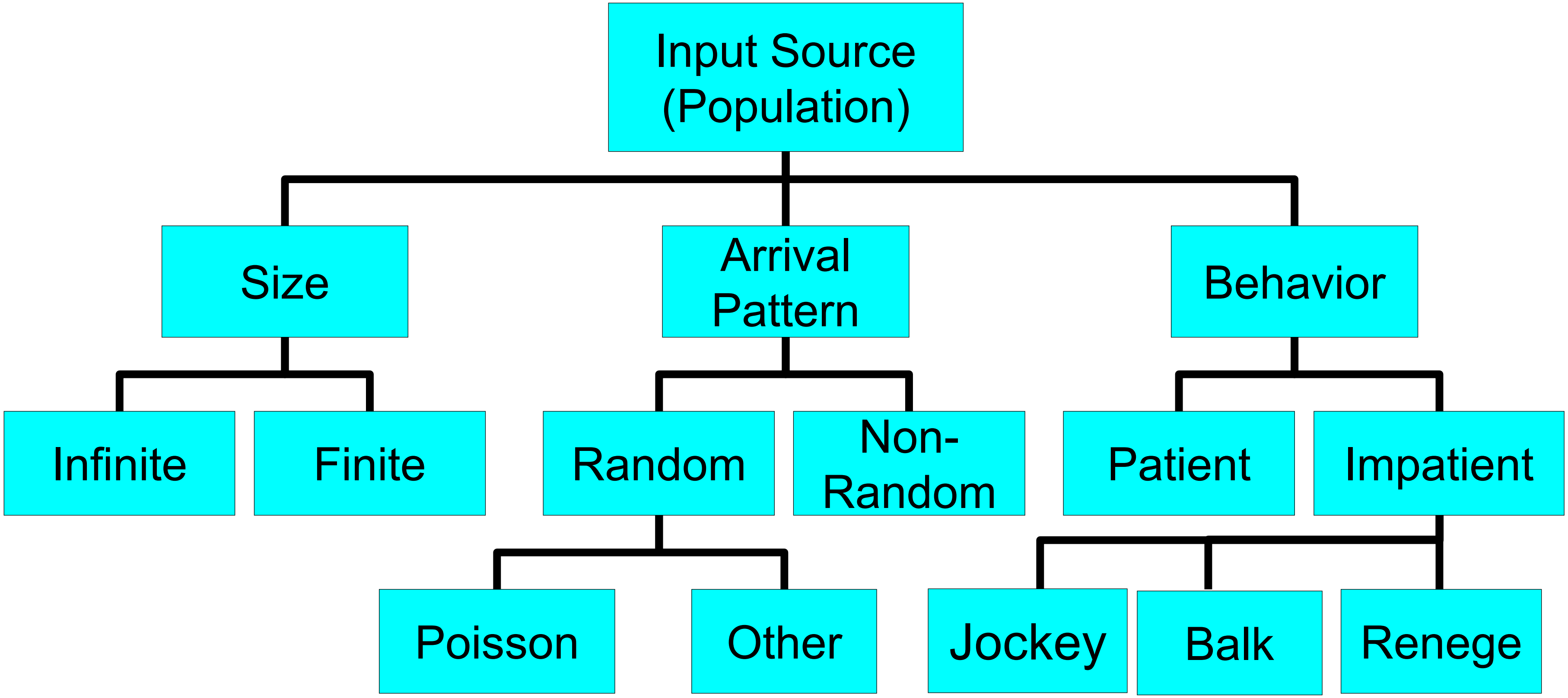
Characteristics of queue components

Waiting Line Terminology

- **Queue**
  - Waiting line
- **Arrival**
  - One person, part, etc. that arrives and demands service
- **Queue discipline**
  - Rules for determining the order that arrivals receive service
- **Phase (stage)**
  - Number of steps in service; after getting service from one queue, move to the next queue for another service.
- **Channel**
  - Number of parallel servers

Arrival Characteristics

- Size of Population
- Arrival Pattern
- Behavior of arrivals



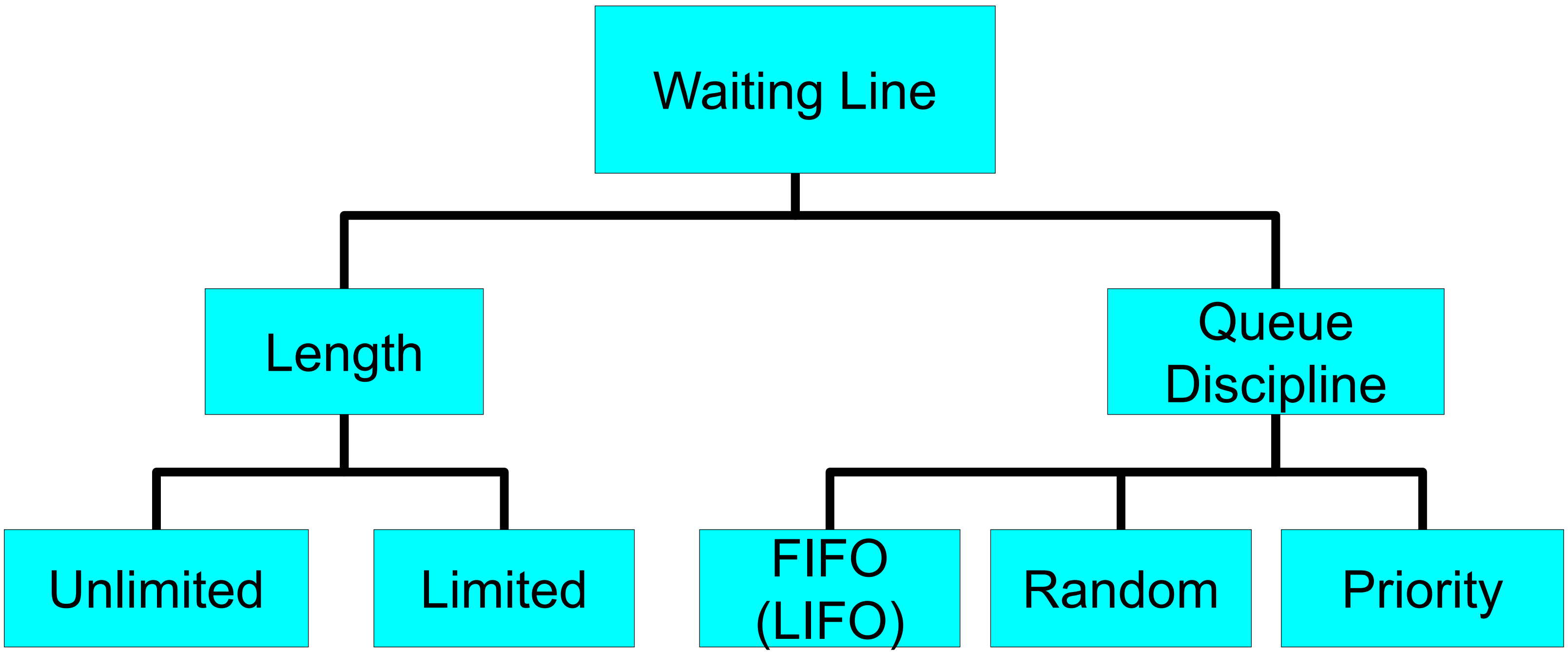
Balk – customers refuse to join the waiting line because it is too long for them

Renege – customers who are already in the waiting line but leave without being served

Jockey – customers switch between waiting lines

Waiting Line Characteristics

- Length of Queue – physical restrictions
- Queue Discipline – rule by which customers in the line are to receive service

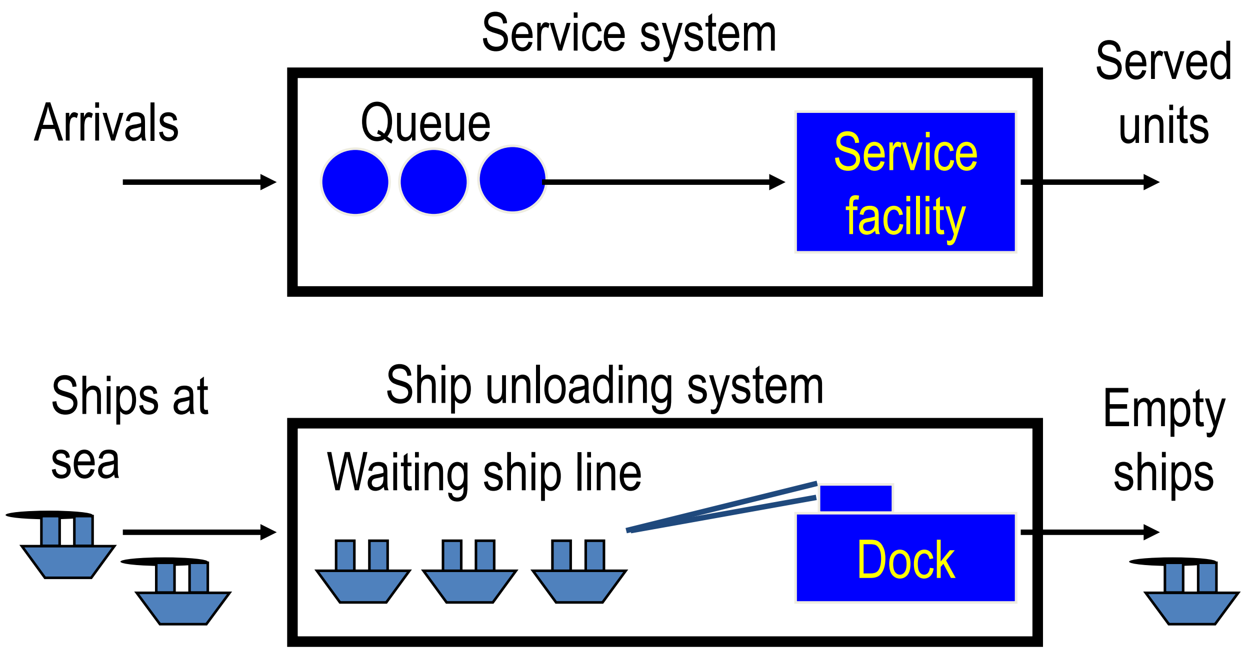


*** FIFO (FCFS) – First-In-First-Out (First-Come-First-Serve)**

*** LIFO – Last-In-First-Out**

Service Characteristics

- Design/configuration of the Service system
- Distribution of Service time Pattern

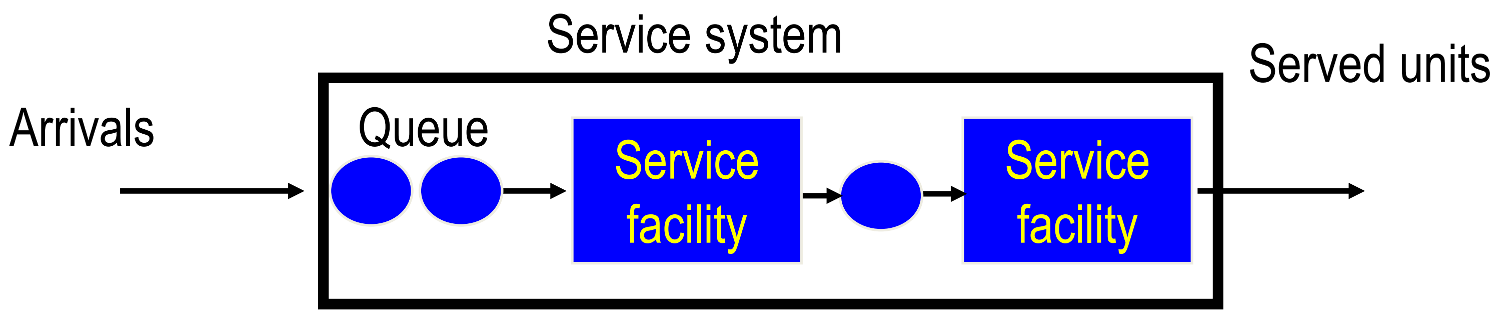


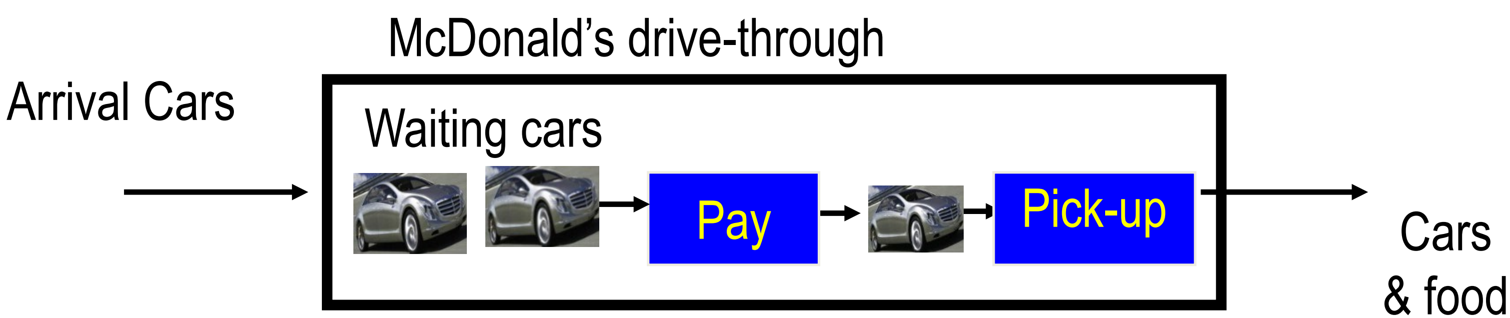
<u>Service System Configuration: Single-Server, Single-Phase System</u>

Examples:

- Customers waiting in line to pay at a small convenient stall with single cashier.
- Customers waiting in line to order and pay at a food stall.

Service Characteristics



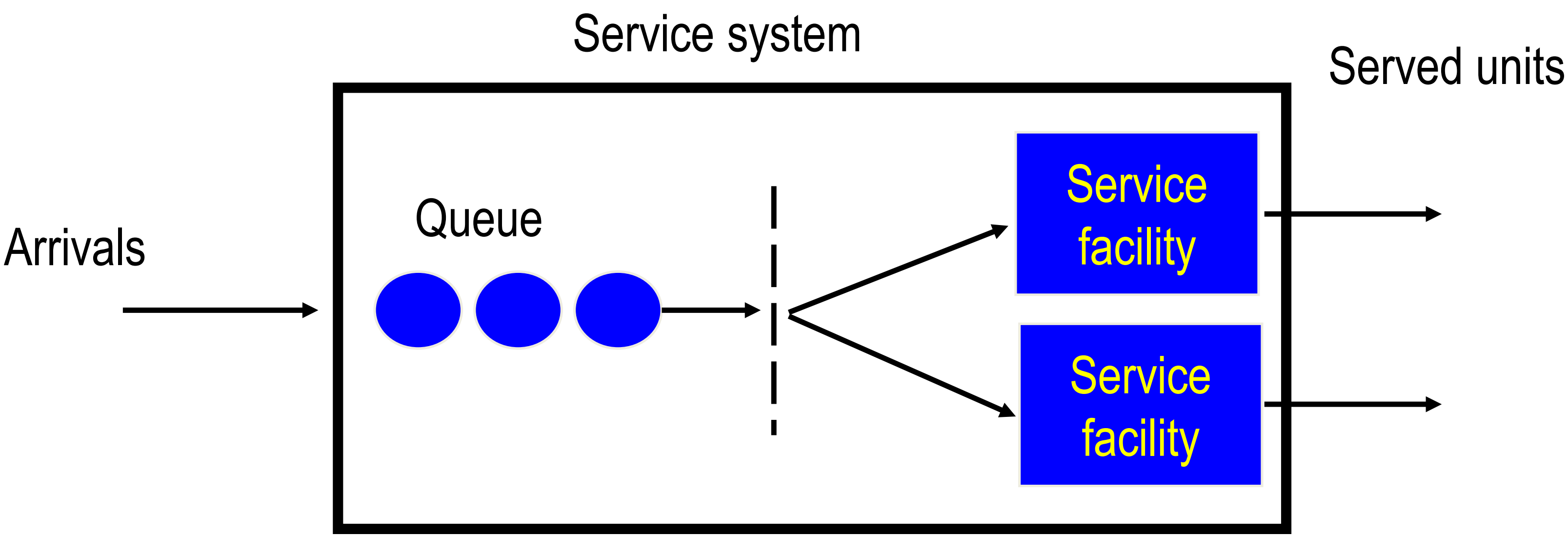


<u>Service System Configuration: Single-Server, Multi-Phase System</u>

Examples:

- Fast food drive-through with different order and pay counters.
- Novelty photo taking booth (single booth) at events with 1 counter for photo taking and 1 counter for collection of print.

Service Characteristics

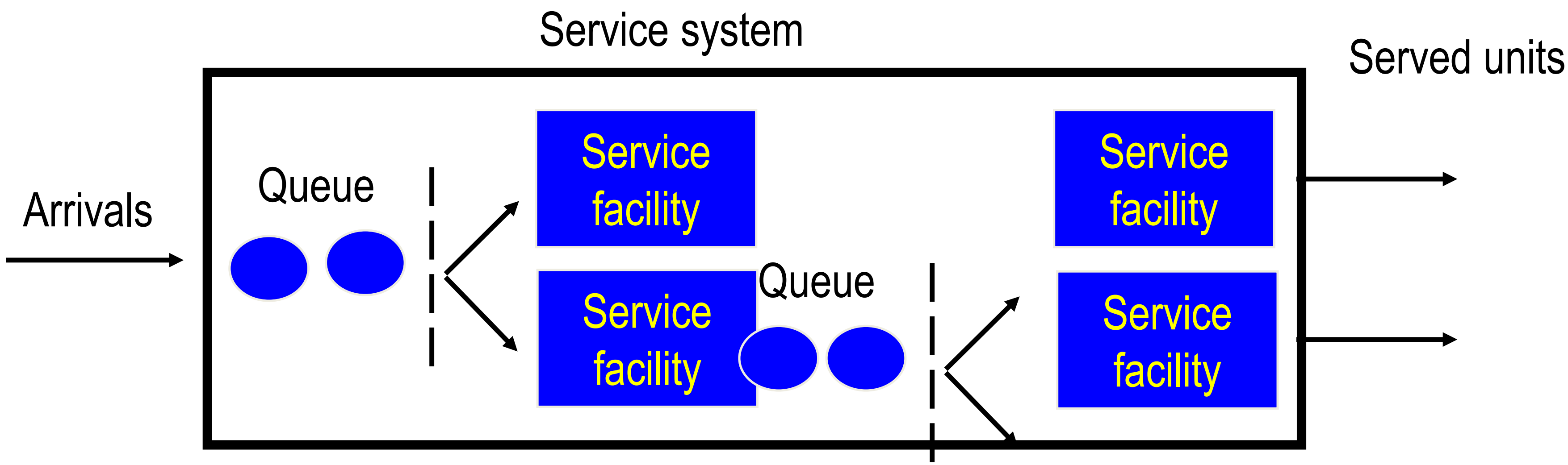


<u>Service System Configuration: Multi-Server, Single Phase System</u>

Examples:

- Bank customers wait in line for one of several tellers.
- Mall customer wait in line for one of the several customer service staff.

Service Characteristics

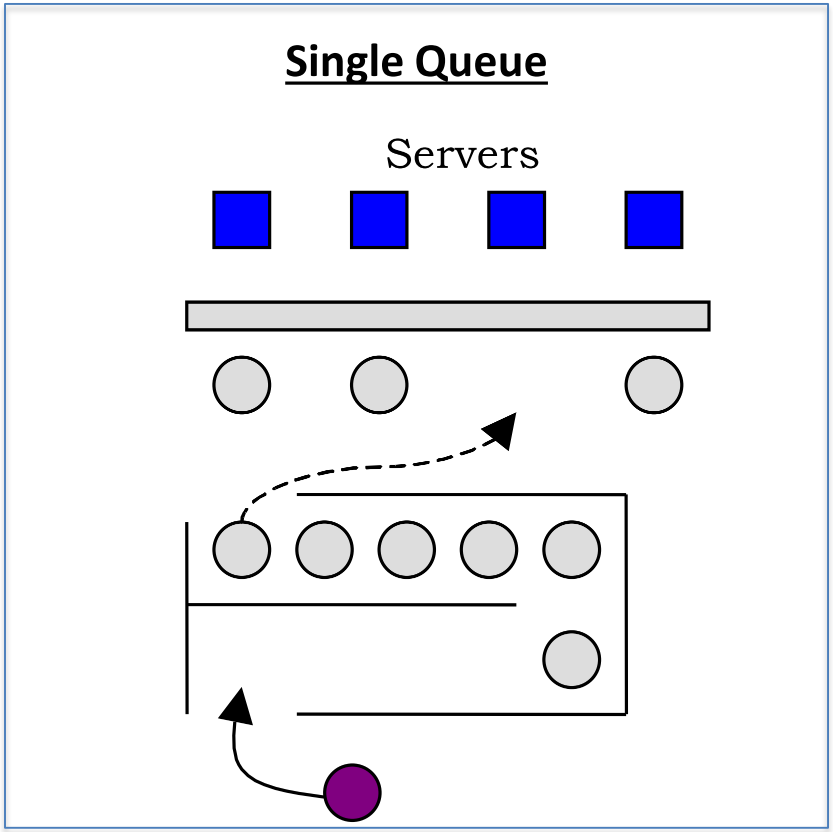


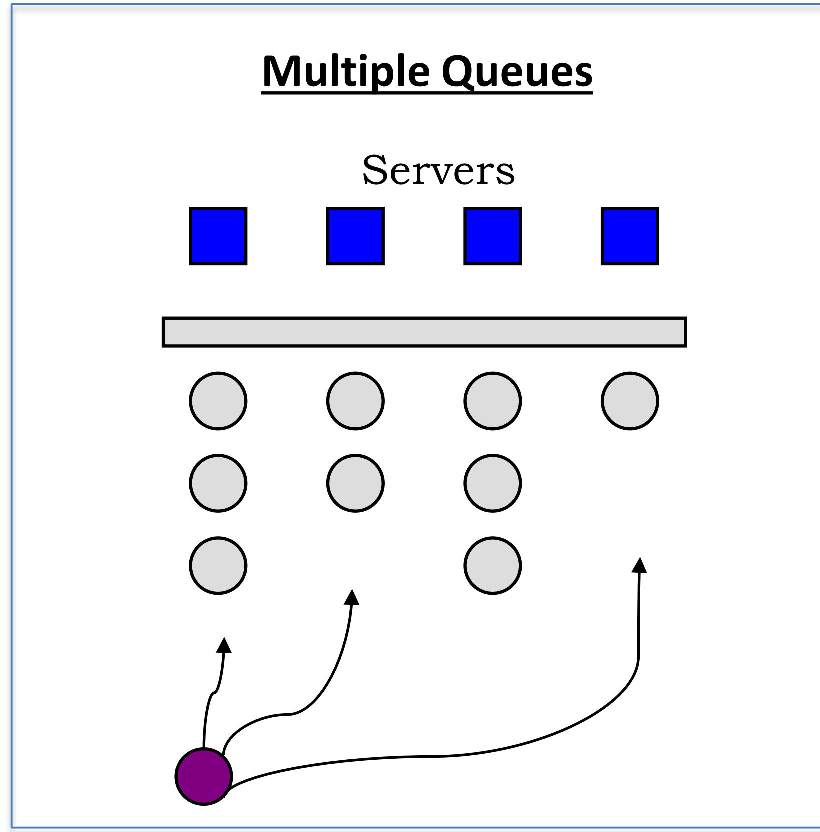
<u>Service System Configuration: Multi-Server, Multi-Phase System</u>

Examples:

- At a laundromat, customers use one of several washers, then one of several dryers.
- Theme park - Queue to buy entrance tickets at the ticket booth, then queue to enter the theme park.

Queue Configurations: Single Queue or Multiple Queues?

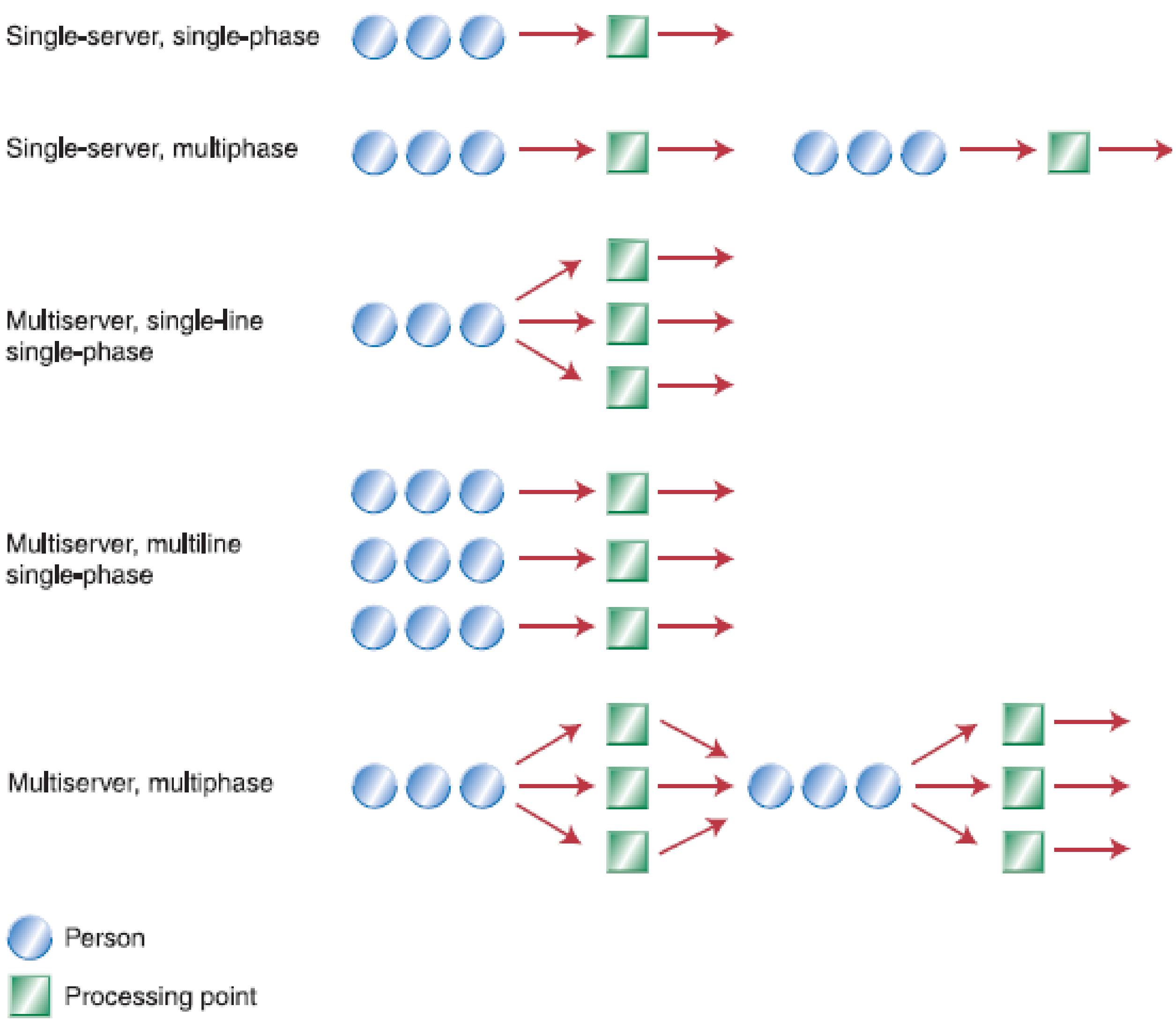




- **Guarantees fairness**
  FIFO applied to all arrivals
- **No customer anxiety regarding choice of queue**
- **Avoids "cutting in" problems**
- **The most efficient set up for minimizing time in the queue**
- **Jockeying (line switching) is avoided**

- **The service provided can be differentiated**
  Example: Supermarket express lanes
- **Labor specialization possible**
- **Customer has more flexibility**
- **Balking behavior may be deterred**
  Several medium-length lines are less intimidating than one very long line

Summary of Waiting Line Systems



http://www.csus.edu/indiv/b/blakeh/mgmt/documents/OPM101SupplC.pdf

Random Effects of Queuing

- Waiting line forms because of **imbalance between demand for service** and the **capacity of the system** to provide the service. Customers usually arrive at unpredictable intervals. The rate of producing service also often varies, depending on customer needs.

- Waiting lines can develop even if the time to process a customer is constant. E.g. train scheduled to arrive every 10 minutes and queues develop while customers wait for the next train or cannot get on a crowded train. Therefore, **variability in the rate of demand** determines the size of waiting lines.

- In general, if there is **no variability** in the demand and service rates and enough capacity has been provided, **no waiting lines** will form.

Matching of the Service Speed

- It may be obvious that the way to keep queues to a minimum is to match the average speed at which people arrive with the average speed at which they are served.

- But this method will definitely create long queues.
  - The interval between arrivals of people is random
  - The service time at the servers is also random
  - Randomness ensures that sometimes there will be no one waiting to be served; and at other times, a whole bunch of people will queue up.

- So even if the service rate is equal to (or higher than) the arrival rate, it is possible that you need to deal with the backlog from previous arrivals.

Given the randomness in arrivals and the service times, how to describe the inter-arrival and service times?

Example: Randomness in Arrival Process

- Assume that the average time between patients' arrivals at a clinic is 15 minutes. How long is the likely time interval until next patient's arrival if
  - 3 minutes have passed since the last patient's arrival?
  - 5 minutes have passed since the last patient's arrival?
  - 11 minutes have passed since the last patient's arrival?
- The likely time interval until next arrival will always be 15 minutes.
- The time elapsed since last arrival does not have any effect on the time interval until next arrival. This is the phenomenon of 'randomness' or 'memoryless' property.
- Memoryless means that the expected time until the next event is the same no matter how long has elapsed since the last event occurred

The randomness in the arrivals shows that:

- You cannot predict the future based on the past if the situation is completely random.
- What probability distributions can be used to describe the arrival process: number of arrivals within a period of time and inter-arrival time?

Poisson Distribution

- Poisson distribution describes the number of events that occur in an interval of time
- It is often used to describe the random arrivals to a queuing system

$$P(X) = \frac{e^{-\lambda}\lambda^{X}}{X!}$$

Where:

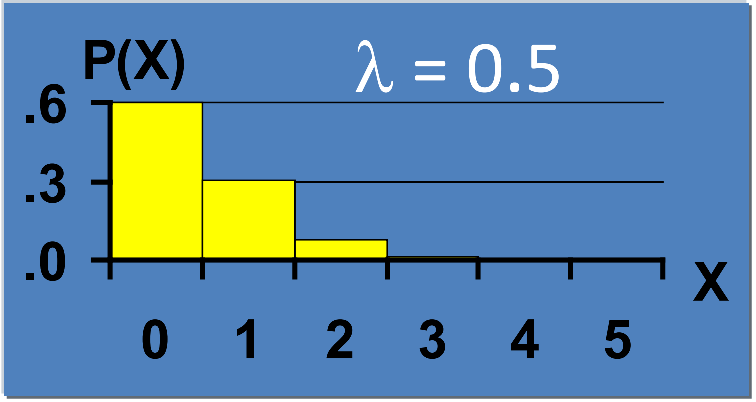
X = no. of arrivals within **given time interval**

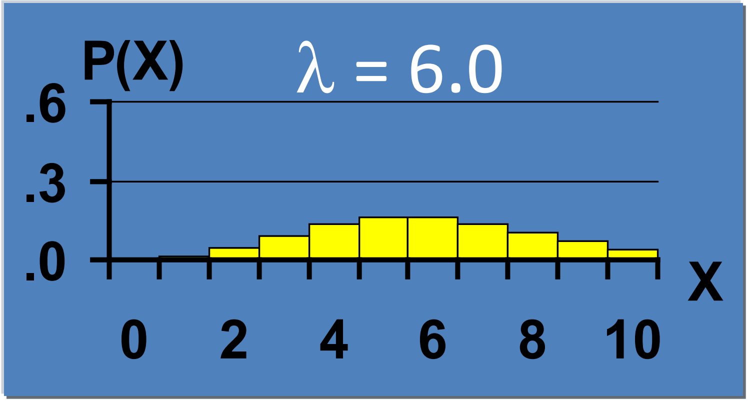
(X = 0,1,2,....)

P(X) = probability of exactly X arrivals **within the time interval**

λ (**Lamda**)= average number of arrivals **within the time interval**

e = 2.7183 (exponential constant)





Will λ for probability of 3 passengers arrived within 3 minutes and within 3 hours remain the same?

Example

A Bank Customer Service Counter serves the needs of customers. Customers arrive at the rate of 8 per hour, on average.

a. What is the probability that no customer arrives at the Customer Service Counter within an hour?

λ = 8 per hour

P(X=0) = $e^{-8} * 8^0 / 0!$= *POISSON.DIST(0,8,0)* **= *0.0003355***

b. What is the probability that three or more customers arrive at the Customer Service Counter within 30 minutes?

λ = 8 per hour = 4 per 30 minutes

P(X>=3) = 1-P(X<3) = 1 – [$e^{-4} * 4^0/0!$ + $e^{-4} * 4^1/1!$ + $e^{-4} * 4^2/2!$]

= 1 – POISSON.DIST(2,4,1) **= *0.7619***

Negative Exponential Distribution

- Negative exponential distribution describes **service time** and **time between arrivals**.

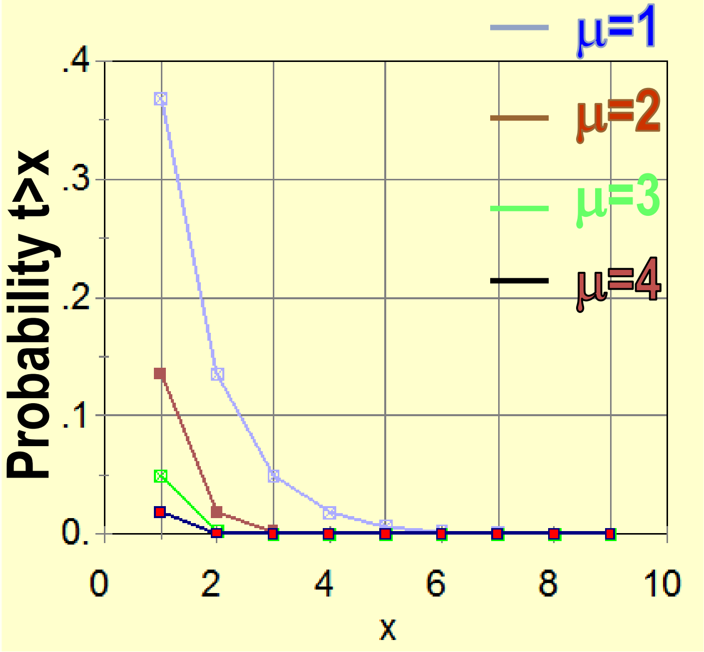
$$P(t) = e^{-\mu t}$$

Where:

t = service time (or time between arrivals) (t >= 0)

P(t) = probability that service time (or time between arrivals) will be greater than t

µ (mu) = average service rate (arrival rate)



How are Poisson and Exponential distributions related?

If customer arrivals follow Poisson distribution , then time between such arrivals (inter-arrival time) follows an exponential distribution.

If service time follows exponential distribution, then number of customers served follows Poisson distribution.

Example

A Bank Customer Service Counter serves the needs of customers. The average time required for servicing is 15 minutes per customer.

a. What is the probability that the Customer Service Counter will need to spend more than 5 minutes to serve a customer?

μ= 1/15 customer per minute

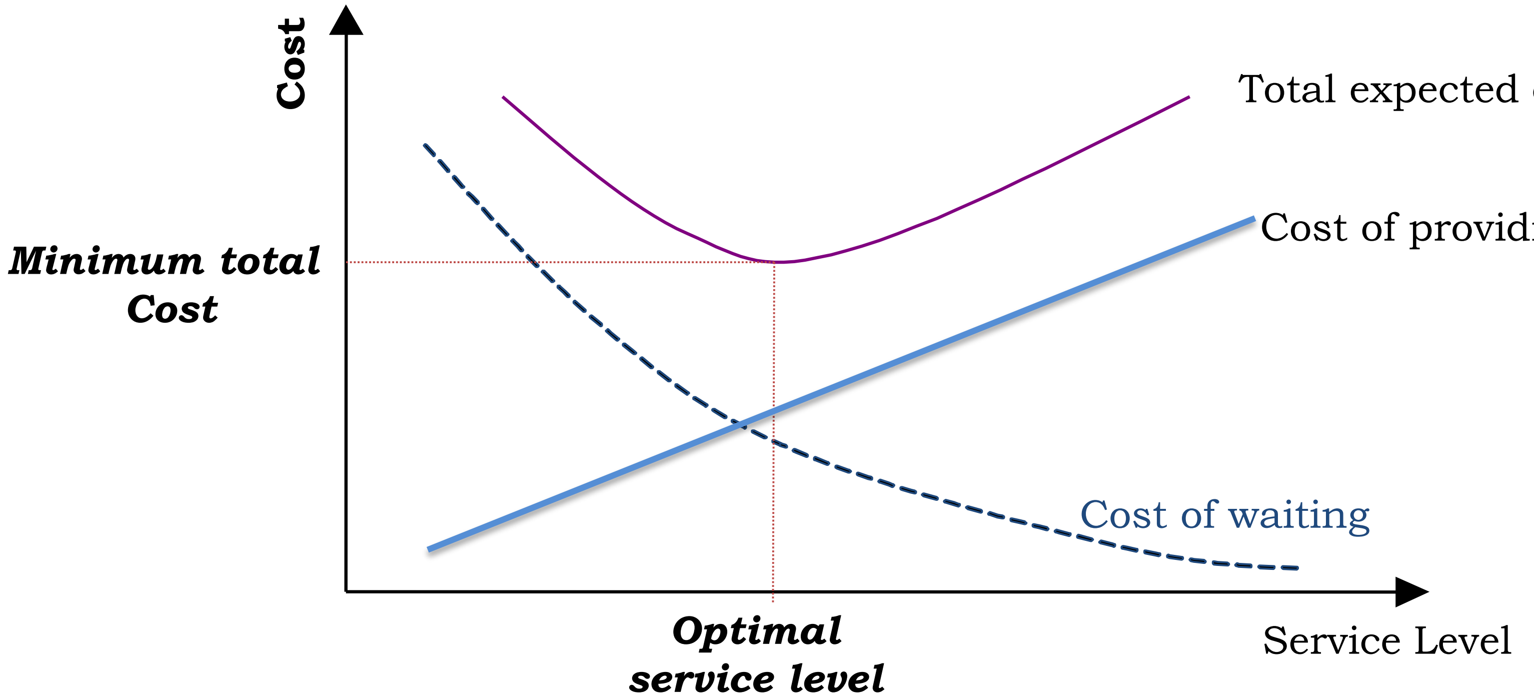
P(t>5) = $e^{-(1/15)*5}$ = *1- EXPON.DIST(5, 1/15, 1)* **= *0.71653***

b. What is the probability that the Customer Service Counter will need to spend 10 minutes or less to serve a customer?

μ = 1/15 customer per minute

P(t<=10) = *1-* $e^{-(1/15)*10}$ = *EXPON.DIST(10, 1/15, 1)* **= *0.48658***

Cost of Waiting in Line

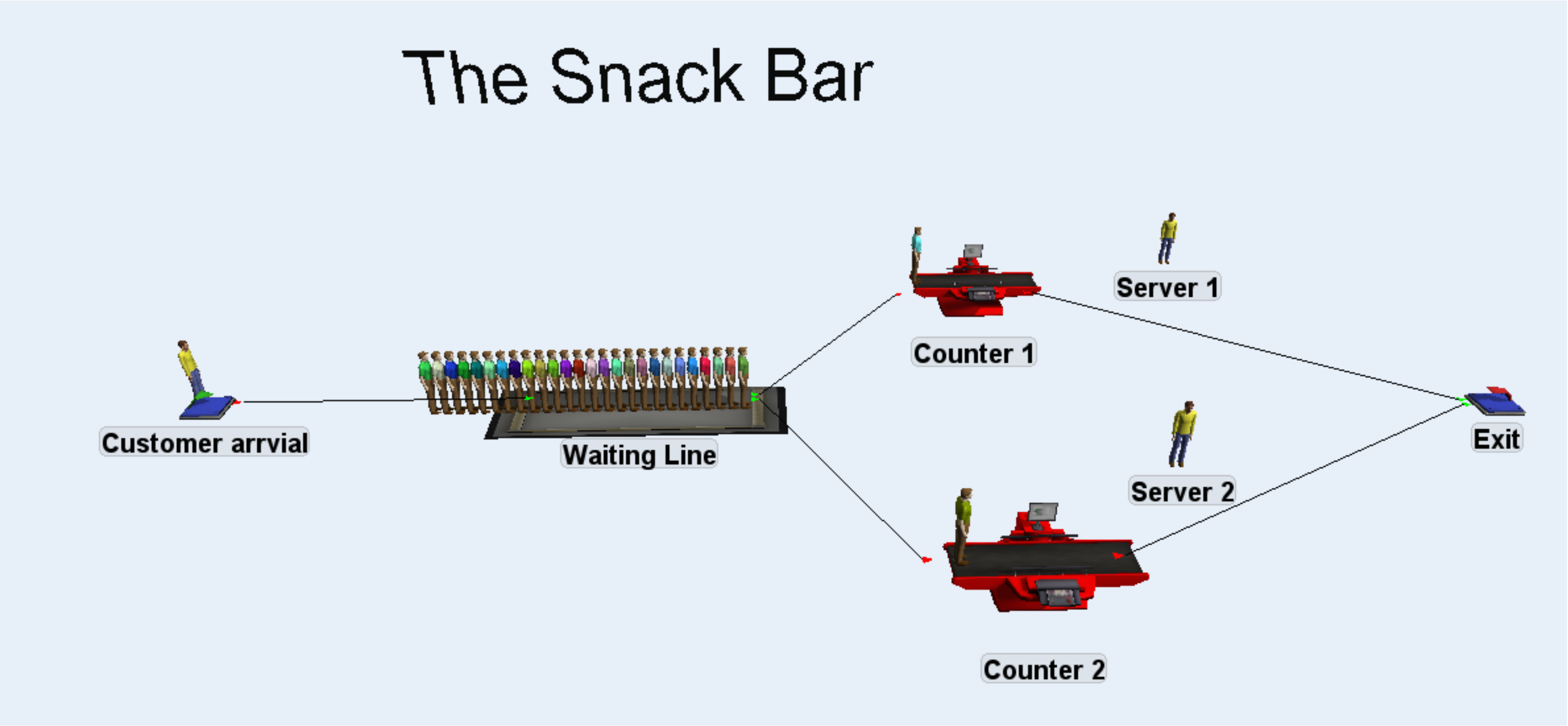


- The relationship between cost of a queuing system and service level can be expressed graphically above.
- Initially, the cost of waiting in line is at a maximum when the organization is at the minimum level of service.
- As level of service increases(e.g. more service counters), there is a reduction in the number of customers in the line and in their waiting times, which decreases queuing cost. However, this will increase the cost of providing the service.
- The optimal service level corresponds to the point with the minimum total cost.

P07 Suggested Solution

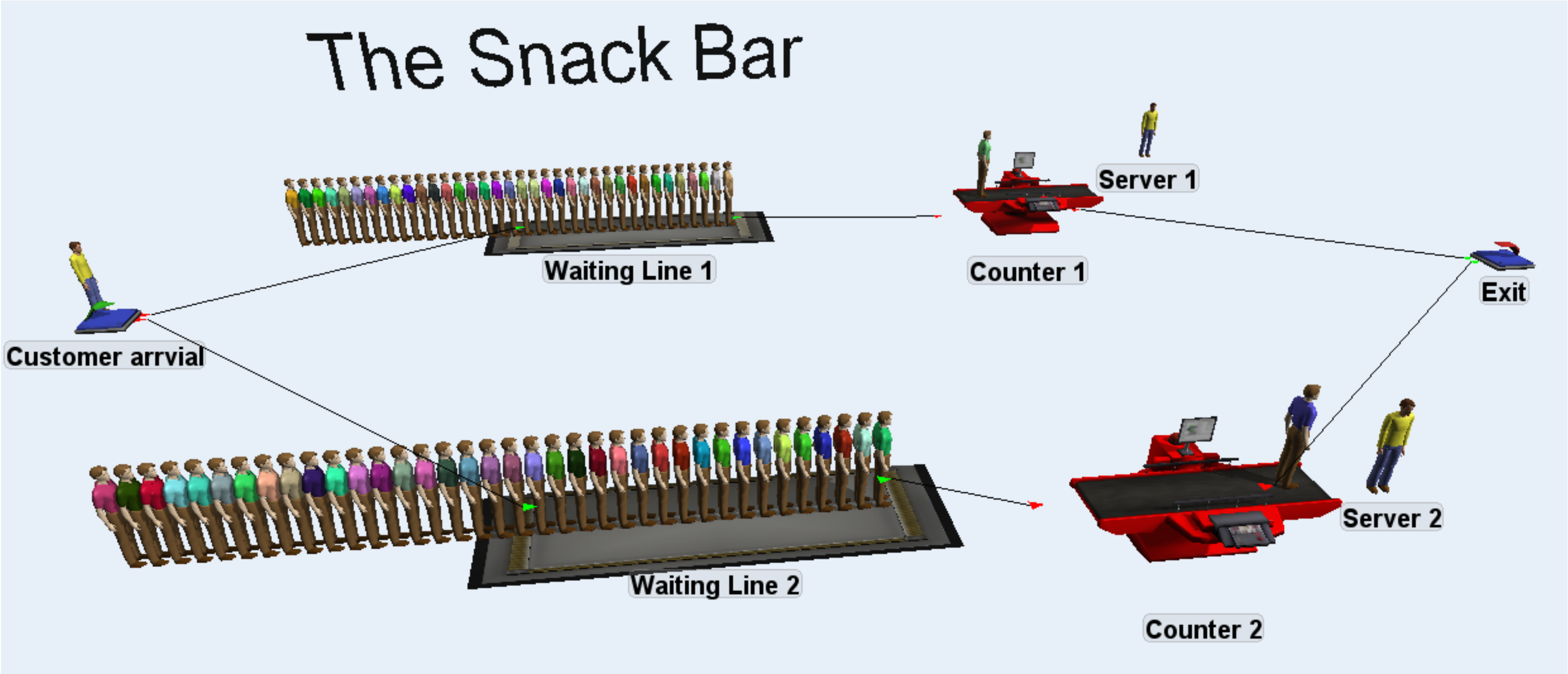
Possible Queuing System

- **Multiple Servers**, Single line



Possible Queuing System

- Multiple Servers, **Multiple lines**



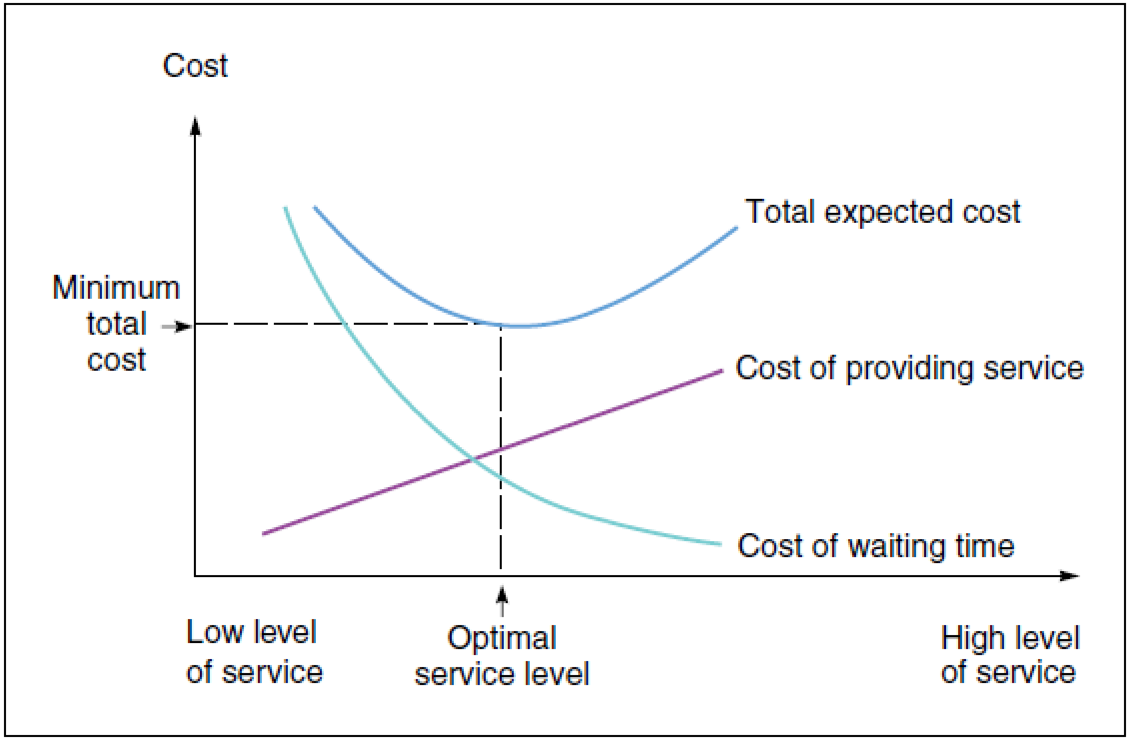
Responses to Lydia's Questions

- **What kind of knowledge is used to address such waiting line problems?**
  - Queuing theory
- **What could have caused the long queue to form?**
  - As observed, customers usually arrive randomly at unpredictable intervals. The rate of producing service also often varies, depending on customers' requirements.
  - Variability in the rate of customers' arrival also determines the size of waiting lines. For example, at times a group of customers come to the snack bar at the same time or no customer for the next 10 minutes.

Responses to Lydia's Questions

➢ **Should the snack bar assign more staff to serve the customers faster?**

- Cost consideration
- Resources (staff) limitations



Responses to Lydia's Questions

How to reduce waiting time and/or queue length?

➢ Increase the Number of Staff and Counters

- Increase the available number of staff at the counters during peak hours to minimize customers' perceived waiting time.
- Opening of more counters will incur higher cost to the snack bar. The snack bar should consider balancing the cost of long customer waiting time and cost associated with opening more counters.

➢ Increase Efficiency of the Service Counters

- Consider additional training for counter staff so as to reduce the service time.
- Have (more) special counters that serve customers who need special orders. Taking orders at the queue to shorten the service time at the counter.
- Allocate one more staff at the counter to speed up the service.

Responses to Lydia's Questions

- **Matching of arrival rate with service rate:**
  - Matching the average speed at which people arrive with the average speed at which they are served will definitely create long queues due to the randomness in arrivals and service times.

- **Queue configuration: with single queue feeding to the respective counters**
  - It can guarantee fairness by avoiding 'cutting in' problems and switching between queues. It is the most efficient set up for minimizing time in the queue.
  - But service provided can not be differentiated. Customers may be intimidated by the long queue and being frustrated in queuing.

Probability Calculations

- Given that customers arrive at the snack bar at an average rate of 1 every 2.5 minutes, the probability that 10 or more customers will arrive in **half an hour**
  - λ = (1/2.5)*30 = 12 customers in half an hour
  - Using Poisson distribution formula:

    $P(X>=10) = 1 – [P(X=0) + P(X=1) + P(X=2) +\ldots\ldots+P(X=9)]$

    $= 1- \sum(\lambda^n/n!)\, e^{-\lambda} = 1-[(12^0/0!)\, e^{-12}+ (12^1/1!)\, e^{-12} + (12^2/2!)\, e^{-12} + \ldots\ldots+ (12^9/9!)\, e^{-12}]$

    = 0.7576

    Using Excel Function:

    P(X>=10) = 1-[POISSON.DIST (0,12,0) + POISSON.DIST (1,12,0) + POISSON.DIST (2,12,0) +… + POISSON(9,12,0)]

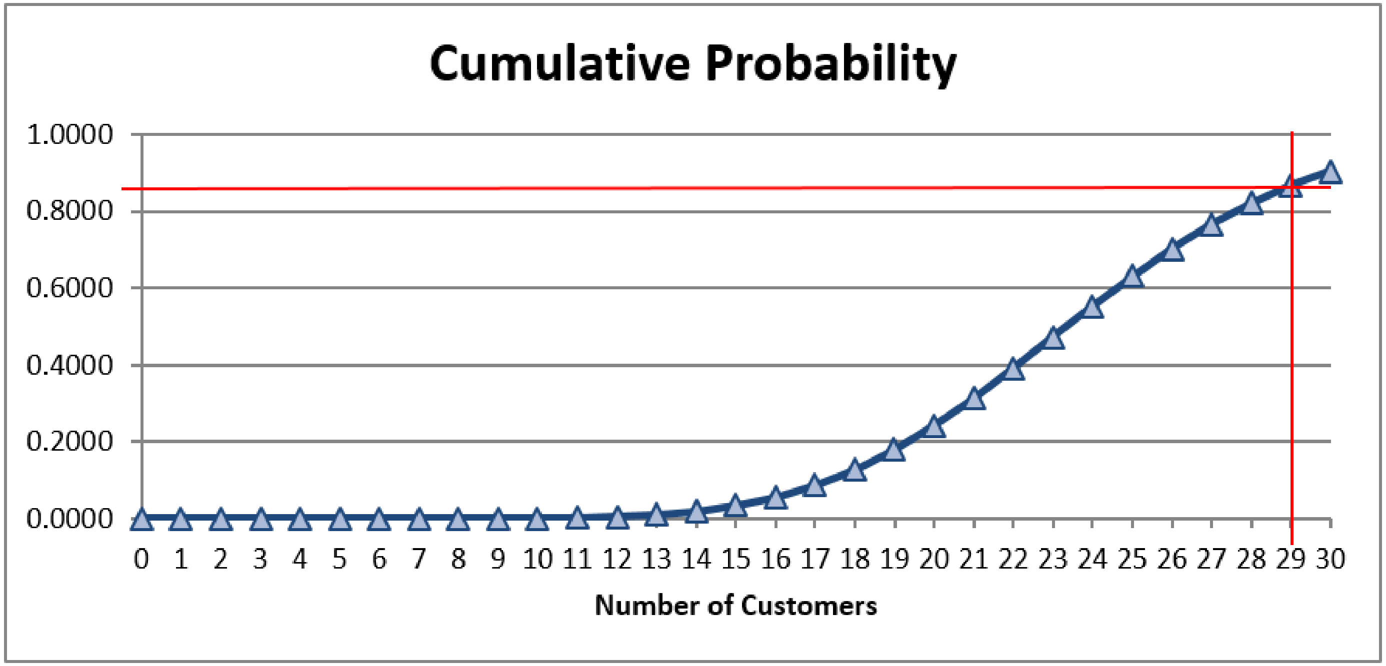
    = 0.7576

    Or P(X>=10)=1- POISSON.DIST(9,12,1) **= 0.7576**

Probability Calculations

- To find the maximum number of customers arrive at any hour at a confidence level of 0.85 with λ = 24 per hour.
- By working out P(X<=K)>=0.85:

K	P(X<=K)
21	0.313928
22	0.391698
23	0.47285
24	0.554001
25	0.631907
26	0.703819
27	0.767742
28	0.822532
29	0.867876



- For at least 85% of the cases, the number of customers arrived at the snack bar per hour will not exceed 29. The snack bar can make use of this information to prepare the amount of food and drinks per hour to meet the customer demand at 85% confidence level.

Probability Calculations

- If service rate of the counter is 1 every 5 minutes (which is the same as 1/5 customers per minute), the probability that one customer will require no more than 3.5 minutes is:

  Using exponential distribution function

  $P(t \leq T) = P(t \leq 3.5) = 1 – e^{-\mu 3.5} = 1 – e^{-(1/5)(3.5)} = 1- e^{-3.5/5}$

  = **0.5034**

  Using Excel Function:

  $P(t \leq T) = P(t \leq 3.5)$ = EXPON.DIST(3.5, 1/5, 1) = **0.5034**

- Having more counters would not change the probabilities but would reduce the queue length and waiting time.

Learning Objectives

At the end of the lesson, students should be able to:

- Explain why a queue forms.
- Identify the three components of a queuing system.
- Describe the basic characteristics of a queuing system and suggest common distributions to represent the random demand and service elements.
- Apply the Poisson / Exponential distribution to model and calculate the random arrival and service characteristics of a queue.
- Appreciate the cost of queuing.

Overview of E211 Operations Planning II Module