Module 2:

Topic 2: D/D/1 and M/M/1 Queuing Models

1. Kendall Notation

1.1. Nation

In queuing theory, Kendall's notation (or sometimes Kendall notation) is the standard system used to describe and classify the queuing model. It was suggested by D. G. Kendall in 1953 as a 3 factor **A/B/s** notation system for characterizing queues.

- A: the arrival distribution
- B: the service (departure) distribution
- s: the number of servers for the system.

1.2. Assumptions

Kendall Notation generally assumes the following:

- Assumption 1: one phase
- Assumption 2: one queue (even with multiple servers)
- Assumption 3: FCFS (First come first served)
- Assumption 4: stable System (Capacity > Arrival rate)
- Assumption 5: “s” is the number of “identical” servers (same capacity for each server, m)

1.3. Symbols

Frequently used symbols for the arrival and service processes are:

- M - Markov distributions (Poisson/exponential)
- D - Deterministic (constant)

For example, M/M/3 refers to a system in which arrivals occur according to a Poisson distribution (i.e., inter-arrival time follows a exponential distribution), service times follow an exponential distribution (i.e., service rate follows a Poisson distribution), and there are 3 servers working at identical service rates.

1.4. Models

We will learn four fundamental queuing models in this learning module.

	Kendall Notation	Arrival Pattern	Service Pattern	Example
1	D/D/1	Constant	Constant	A raw material is entering an automated assembly line exactly every 5 min and operation takes exactly 4 min.
2	M/M/1	Exponential	Exponential	Students arrive at a copier in library at an (average) rate of 15/hr. Each spends an average of 3 min. making copies.
3	M/D/1	Exponential	Constant	An automated pizza vending machine heats and dispenses a slice of pizza in exactly 3 minutes. Customers arrive at an average rate of one every 4 minutes.
4	M/M/s	Exponential	Exponential	There are two Drive-up windows at a fast food restaurant. Customers arrive at the average rate of 24 per hour. Each employee can serve one customer every two minutes on average.

1.5. Performance Measurements in Queuing System

Notation	Description
n_s	Average number of customers in the system(waiting and being served) NOTE: It can be interpreted as the work in process (WIP) of the queuing system.
n_l	Average number of customers in the queue (line) NOTE: It can be interpreted as the work in process (WIP) ONLY in the waiting line.
t_s	Average time a customer spends in the system (waiting and being served) NOTE: It can be interpreted as the throughput time (TP) of the queuing system.
t_l	Average time a customer spends waiting in queue (line) NOTE: It can be interpreted as the throughput time (TP) ONLY in the waiting line.
P_0	Probability of no (zero) customers in the system
P_n	Probability of n customers in the system
ρ (Rho)	Server Utilization rate (i.e., % of time that the server is busy) NOTE: It can be interpreted as the resource utilization of the queuing system.

2. Model 1: D/D/1

Example 1: D/D/1

A raw material is entering an automated assembly line exactly every 5 min and operation takes exactly 4 min.

a. Arrival rate = ?

b. Service rate = ?

c. What % of time is the server (machine) busy?

d. Avg time in queue?

e. Avg time in system?

f. Avg number of materials in queue?

g. Departure rate or Output rate?

h. Average number of materials in system?

3. Model 2: M/M1

3.1 Formula

When the arrival or service pattern is not deterministic (i.e., constant assumption) it is not simple derive the above performance measurements. Since this is an introductory course, we will simply apply the formula given below instead of deriving each one.

Probability that server is busy, server utilization	$\rho = \dfrac{\lambda}{\mu}$
Probability that no customers are in system	$P_0 = \left(1 - \dfrac{\lambda}{\mu}\right)$
Probability of exactly n customers in system	$P_n = \left(1 - \dfrac{\lambda}{\mu}\right)\left(\dfrac{\lambda}{\mu}\right)^n$
Average number of customers in system	$n_s = \dfrac{\lambda}{\mu - \lambda}$
Average number of customers in queue (line)	$n_l = \dfrac{\lambda^2}{\mu(\mu - \lambda)}$
Average time customer spends in system	$t_s = \dfrac{1}{(\mu - \lambda)}$
Average time customer spends in queue (line)	$t_l = \dfrac{\lambda}{\mu(\mu - \lambda)}$

3.2. Example

Example 2: M/M/1
Students arrive at a copier in library at an (average) rate of 15/hr. Each spends an average of 3 min. making copies.
a. Copier utilization?
b. Avg time in queue?
c. Avg no. of customers in queue?
d. Avg no of customers in system?
e. (Average) Departure rate?
f. Probability of exactly 1 customer in queue?

Example 3: M/M/1 Costing Problem
Calls arrive for a single customer service rep (CSR) at a rate of 20/hr and take an average 2.5 minutes to process. If the CSR is busy, the caller is put on hold.
The CSR is paid $20/hr. It has been estimated that each minute a customer spends on hold costs $2 due to customer dissatisfaction and loss of future business.
Estimate the average hourly cost, assuming that it is a M/M/1 queuing system.