UNIT - 2

Queuing Theory

* Queuing theory :- A group of items waiting to receive services, including those receiving the service is known as waiting line or a queue.
Waiting line or queue are omnipresent. Business of all types, industry, school, hospitals, school, hospitals, book stalls, banks, post office, petrol pump, all have queuing problems.

* The basic queuing process and its characteristics -
The basic queuing process can be described as a process in which the customers arrive for service at a service counter or station, wait for their turn in the queue if the server is busy in the service of the customer and are served when the server gets free and the customer leave the system as soon as he is served.

Characteristics of Queuing System -

1- Arrival distribution - It represents the pattern in which the number of customers arrive at the system. Arrivals may also be represented by the inter-arrival time, which is the period between two successive arrivals.
The rate at which customers arrive to be serviced, i.e. number of customers arriving per unit of time is called arrival rate. When the arrival rate is random, the customers arrive in no logical pattern.

This represents most cases in the business world.

2- Service distribution – It represents the pattern in which the number of customers leave the system. Departures may also be represented by the service time which is the time between two successive services.

3- Service channels – The queuing system may have a single service channel. Arriving customers may form one line and get serviced, as in a doctor's clinic. In case of parallel channels, several customers may be serviced simultaneously as in a barber shop. A queuing model is called one service model when the system has one server only and a multi-server channel (model) when the system has a number of parallel channels each with one server.

4- Service discipline – The service discipline refers to the manner in which the members in the queue are chosen for service. The following service disciplines are seen in common practice.

(i) First Come, First Serve (FCFS) – According to this discipline the customers are served in the order of their arrival. This service discipline may be seen at a railway ticket window etc.

(ii) Last Come, First Serve (LCFS) – According to this discipline the items arriving last are taken out first. This discipline may be seen in godowns where the units (items) which come last are taken out first.

(iii) Service in random order.

(iv) Service on some Priority – Procedure – Some customers are served before the others without considering their order of arrival.

5- Maximum number of Customers allowed in the system – The Maximum numbers of customers in the system can be either finite or infinite. In some facilities, only a limited number of customers are allowed in the system and new arriving customers are not allowed to join the system.

6- Service Mechanism – The service mechanism refers to
(i) the pattern according to which the customers are served.
(ii) facilities given to the customers:

(a) Single-channel: Here the customers are served by one counter only.

(b) Multi-channel: Here the customers are served by more than one counter.

★ Definitions in Queuing Problems :-

1- Queue length – Queue length is defined by the number of persons (customers) waiting in the line at any time.

2- Average length of line – Average length of line is defined by the number of customers in the queue per unit time.

3- Waiting time – It is time upto which a unit has to wait in the queue before it is taken into service.

4- Servicing time – The time taken for servicing of a unit is called its servicing time.

5- Idle period – When all the units in the queue are served. The idle period of the server begins and it continuous upto the time of arrival of the unit. The idle period of a server is the time during which he remains free because there is no customers present in the system.

6- Mean arrival rate – The mean arrival rate in a waiting-line situation is defined as the expected number of arrivals occuring in a time interval of length unity.

7- Mean servicing rate – The mean servicing rate for a particular servicing station is defined as the expected number of services completed in a time interval of length unity, given that the servicing is going on throughout the entire time unit.

8- Steady state - A system is said to be in steady state when its operating characteristics becomes independent of time.

9- Traffic intensity - In case of a simple queue the traffic intensity is the ratio of mean arrival rate and the mean servicing rate.

$$\text{Traffic intensity} = \frac{\text{Mean arrival rate}}{\text{Mean servicing rate}}$$

* Poisson Process :-

Theorem :- In Poisson process the probability of n arrivals during time intervals of length t is given by -

$$P_n(\Delta t) = \frac{(\lambda \Delta t)^n}{\lfloor n} e^{-\lambda \Delta t}$$

where λ, t is the parameter.

Proof :-

Case I - $n = 0$ (no arrival)

$$P_0(\Delta t) = \frac{(\lambda \Delta t)^0 e^{-\lambda \Delta t}}{\lfloor 0}$$

$$P_0(\Delta t) = e^{-\lambda \Delta t}$$

$$P_0(\Delta t) = 1 - \lambda \Delta t + \frac{(\lambda \Delta t)^2}{2} - \frac{(\lambda \Delta t)^3}{\lfloor 3} \quad \left[e^{-x} = 1 - x + \frac{x^2}{\lfloor 2} - \frac{x^3}{\lfloor 3} \dots \right]$$

$o(\Delta t) \rightarrow 0$ when Δt is very small

$$\boxed{P_0(\Delta t) = 1 - \lambda \Delta t}$$

Case II — $n = 1$

$$P_1(\Delta t) = \frac{(\lambda\Delta t)^1 e^{-\lambda\Delta t}}{\lfloor 1}$$

$$= \lambda\Delta t \cancel{*} \left[1 - \lambda\Delta t + \frac{(\lambda\Delta t)^2}{\lfloor 2} + \ldots\ldots\right]$$

$$= \lambda\Delta t - (\lambda\Delta t)^2 + \frac{(\lambda\Delta t)^3}{\lfloor 2} + \ldots\ldots$$

$O(\Delta t) \to 0$

$$\boxed{P_1(\Delta t) = \lambda\Delta t}$$

Case III — $n = m > 1$ (m is more than one)

$$P_m(\Delta t) = \frac{(\lambda\Delta t)^m e^{-\lambda\Delta t}}{\lfloor m}$$

$$P_m(\Delta t) = \frac{(\lambda\Delta t)^m}{\lfloor m}\left[1 - \lambda\Delta t + \frac{(\lambda\Delta t)^2}{\lfloor 2} - \frac{(\lambda\Delta t)^3}{\lfloor 3} + \ldots\ldots\right]$$

$\Delta t \to 0$

$$\boxed{P_m(\Delta t) = 0}$$

$$\sum_{n=0}^{\infty} P_n = 1 \Rightarrow P_0 + P_1 + P_2 + \ldots\ldots$$

$$= 1 - \lambda\cancel{\Delta t} + \lambda\cancel{\Delta t} + \ldots$$

$$\boxed{\sum_{n=0}^{\infty} P_n = 1}$$

★ **Model I (M/M/1):(∞/FCFS)(Birth and death model):-**

This is the queuing model with poisson arrival, poisson service, single channel with infinite capacity. The service discipline is first come first service. Here λ is the mean arrival rate and μ is the mean service rate.

To find the steady state equation—

t	arrival capacity (λ) (Δt)	Service capacity (μ) (Δt)	$t+\Delta t$
n	1	1	n
$n+1$	0	1	n
$n-1$	1	0	n
n	0	0	n

$$P_n(t+\Delta t) = P(\text{case I}) + P(\text{case II}) + P(\text{case III}) + P(\text{case IV})$$

$$P_n(t+\Delta t) = P_n(t)(\lambda\Delta t)(\mu\Delta t) + P_{n+1}(t)(1-\lambda\Delta t)(\mu\Delta t) + P_{n-1}(t)\lambda\Delta t(1-\mu\Delta t) + P_n(t)(1-\lambda\Delta t)(1-\mu\Delta t)$$

$$P_n(t+\Delta t) = P_n(t)\lambda\mu(\Delta t)^2 + P_{n+1}(t)(\mu-\lambda\mu\Delta t)\Delta t + P_{n-1}(t)(\lambda-\lambda\mu\Delta t)\Delta t + P_n(t)[1-\lambda\Delta t-\mu\Delta t+\lambda\mu(\Delta t)^2]$$

$$P_n(t+\Delta t) = P_n(t)\lambda\mu(\Delta t)^2 + \mu P_{n+1}(t)\Delta t - P_{n+1}(t)\lambda\mu(\Delta t)^2 + P_{n-1}(t)\lambda\Delta t - P_{n-1}(t)\lambda\mu(\Delta t)^2 + P_n(t)[1-\mu\Delta t-\lambda\Delta t+\lambda\mu(\Delta t)^2]$$

$$P_n(t+\Delta t) = P_n(t)\lambda\mu(\Delta t)^2 + P_{n+1}(t)\mu\Delta t - P_{n+1}(t)\mu\Delta t - P_{n+1}(t)\lambda\mu(\Delta t)^2 + P_{n-1}(t)\lambda\Delta t - P_{n-1}(t)\lambda\mu(\Delta t)^2 + P_n(t) - P_n(t)\mu\Delta t - P_n(t)\lambda\Delta t + P_n(t)\lambda\mu(\Delta t)^2$$

$$\frac{P_n(t+\Delta t)-P_n(t)}{\Delta t} = P_n(t)\lambda\mu\Delta t + P_{n+1}(t)\mu - P_{n+1}(t)\lambda\mu\Delta t + P_{n-1}(t)\lambda - P_{n-1}(t)\lambda\mu\Delta t - P_n(t)\mu - P_n(t)\lambda + P_n(t)\lambda\mu\Delta t$$

$\Delta t \to 0$ (Directional derivative zero)

$$0 = P_{n+1}(t)\mu + P_{n-1}(t)\lambda - P_n(t)\mu - P_n(t)\lambda$$

$$\boxed{P_n(t)(\lambda+\mu) = P_{n+1}(t)\mu + P_{n-1}(t)\lambda}$$

This is steady state equation.

1- Probability of no customers in the system is -

$$P_0 = 1 - \rho \qquad \text{where } \rho = \frac{\lambda}{\mu}$$

2- Probability of n customers in the system -

$$P_n = (1-\rho)\,\rho^n$$

3- Probability of more than n customers in the system

$$P(>n) = \rho^{n+1}$$

4- Probability of more than n customers in the queue

$$P(>n+1) = \rho^{n+2}$$

5- The average (expected) number of customers in the system

$$L \text{ or } L_s = \frac{\lambda}{\mu-\lambda} = \frac{\rho}{1-\rho}$$

6- Average (expected) queue length is

$$L_q = L_s - \frac{\lambda}{\mu} = \frac{\rho^2}{1-\rho}$$

7- Average (expected) waiting time of customer in the queue

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

8- Average (expected) waiting time that a customer spend in the system is

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

9- The variance of queue length

$$Var(n) = \frac{\rho}{(1-\rho)^2}$$

Inter-relationship between L_s, L_q, W_s, W_q -

$$L_s = \lambda W_s$$

$$L_q = \lambda W_q$$

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

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

<u>Ques</u> Customers arrive at a sales counter managed by single person according to a poisson process with a mean rate of 20 per hour. The time require to serve a customer has an exponential distribution with a mean of 100 second. Find the average time of a customer.

<u>Sol:-</u>

$$\frac{1}{\mu} = \frac{10\cancel{0}}{6\cancel{0} \times 6\cancel{0}} = \frac{1}{36}$$

$$\mu = 36 \text{ / hour}$$

$$\lambda = 20 \text{ / hour}$$

Average waiting time in the queue -

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

$$= \frac{20}{36(36-20)} = \frac{20}{36 \times 16} = \frac{5}{144} \times 60 \times 60$$

$$= 125 \text{ / hr}$$

Average waiting time in the system -

$$W_s = \frac{1}{\mu-\lambda} = \frac{1}{36-20}$$

$$= \frac{1}{16} \times 60 \times 60 = 225 \text{ / hr.}$$

Ans.

★ Model II $(M/M/1):(M/FCFS)$ Finite queue length model –

1- The probability of no customer in the system

$$P_0 = \frac{1-\rho}{1-\rho^{N+1}}, \quad \text{where } \rho = \frac{\lambda}{\mu}$$

$\lambda \rightarrow$ mean arrival rate

$\mu \rightarrow$ mean service rate

2- The probability of n customers in the system

$$P_n = \frac{1-\rho}{1-\rho^{N+1}} \rho^n$$

3- Average number of customers in the ~~queue~~ system

$$L_s = \frac{\rho\left[1-(1+N)\rho^N + N\rho^{N-1}\right]}{(1-\rho)(1-\rho^{N+1})} \quad \text{where } \rho = \frac{\lambda}{\mu}$$

4- Average number of customers in the queue

$$L_q = \frac{1-N\rho^{N-1} + (N-1)\rho^N}{(1-\rho)(1-\rho^{N+1})}\rho^2 = \sum_{n=0}^{N} nP_n$$

5- Average time a customer spend in the system

$$W_s = \frac{L_s}{\lambda'} \quad \text{where } \lambda' = \lambda(1-\rho^N)$$

6- Average waiting time in the queue

$$W_q = \frac{L_q}{\lambda'} \quad \text{where } \lambda' = \lambda(1-\rho^N)$$