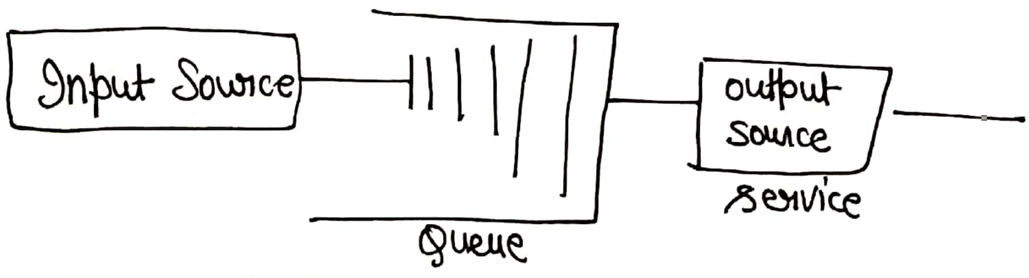
Optimization Techniques

MGKV BCA 4 Sem

Queueing Theory

Queue :- A group of items waiting to receive service are known as queue. Queus contain customers (or items) such as people, objects or information. Queue form when there are limited resources for providing a service i.e. when demand for a service is more than the capacity of service facility then queue created.

Example - waiting in line at a bank or a store.



So, Queuing theory is the mathematical study of queuing or waiting lines. Queuing theory provides a large number of alternative mathematical models for describing and solving waiting line problems.

Hence, the objective to study queuing theory is to maintain a balance between waiting time and service cost so that these two could be minimized.

Characteristics of Queuing System :-

The important characteristics required to study the waiting line are –

1) **Queue Length :–** The average number of customers in the queue waiting to get service.

Queue length ↑ then service performance ↓ and vice – versa.

2) **System Length :–** system length is equal to the average number of customer waiting in the queue and the customer those being served.

3) **Waiting time in the Queue :–** The average time that a customer has to wait in the queue to get service.

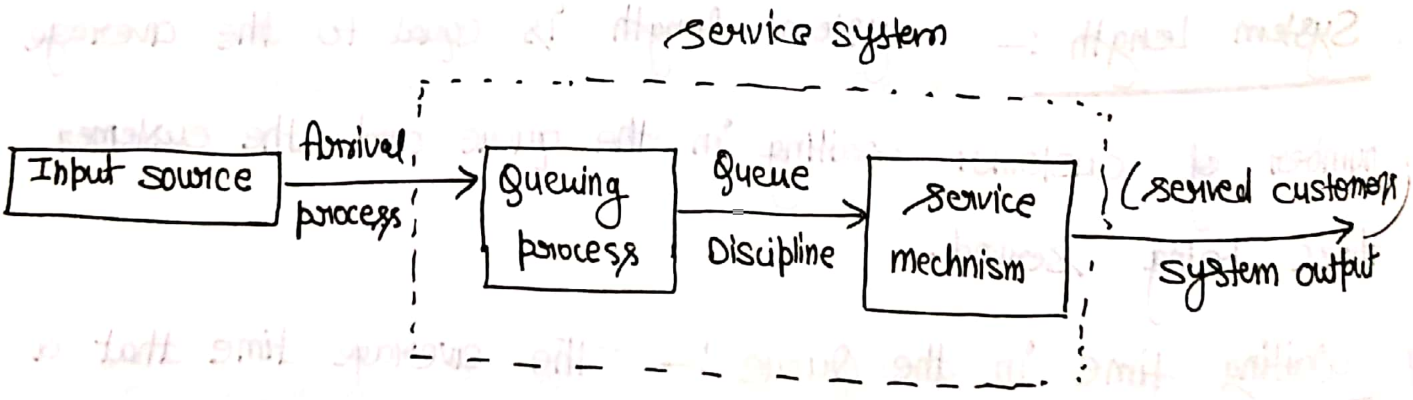
waiting time ↑ then customer dissatisfaction ↑

4) **Waiting Time in the system** :— It implies the time spent by the customer in the queue system (i.e. form entry in the queue to completion of service)

i.e. waiting time in the system = waiting time in queue + service time.

5) **Server Idle Time** :— The average time in which server is not busy to provide any service facility.

Elements of Queuing System :— The various elements of queuing system are defined as—



i.e. The 'customers' requiring service are generated over time by Input Source. These customers enter the queuing system and join the queue. At certain times a member of the queue is selected for service by some rule known as service discipline and then leaves the queue.

Discuss one by one elements as—

1) Input Source / Calling population / Input process / Arrival process:

The Calling population is the source of the customers to the waiting line system. Two important characteristics of the input source are –

(i) Its size (ii) the pattern of arrival.

The size is the number of customer that might require service from time to time. The size of calling population may be finite or infinite.

Infinite population represents a large number of customers that it is always possible for one more customer to arrive to be served.

For example – a departmental store, railway reservation, a bank, tax/toll etc.

whereas a finite population has a specific, countable number of potential customers.

for example – Number of process run by a computer or a certain number of machines repaired by a service man.

The pattern of arrival defines the way customers enter the system. Most arrivals in a service system are at random.

The exact prediction of any arrival in random system is not possible. Therefore, the number of arrivals per unit time is estimated by ~~poisson~~ poisson distribution.

So, mean arrival rate = average number of arrivals per unit time = λ

mean arrival time = time interval between the consecutive arrivals = $\frac{1}{\lambda}$

2) **Queuing Process :-** Queue represent a certain number of customers waiting for service (being served is not considered in the queue.)

The queuing process refers to the number of queues and their respective lengths. The number of queue is single, multiple or priority queue depend upon the layout of a service system.

3) **Queuing Discipline :-** The queue discipline specifies the manner in which the customers waiting in the queue are selected for the service. The rules governing order of service may be –

(i) First In first out (FIFO) — This is also known as first come- first served (FCFS) service discipline. According to FIFO, the customers are served in the order of their arrival.

for example — Bank counters, railway reservation etc.

(ii) Last In first out (LIFO) — According to LIFO, the items arriving last are taken out first.

for example — In government offices.

(iii) Service in random order (SIRO) — under this rule customers are selected for service at random, irrespective of their arrivals in the service system.

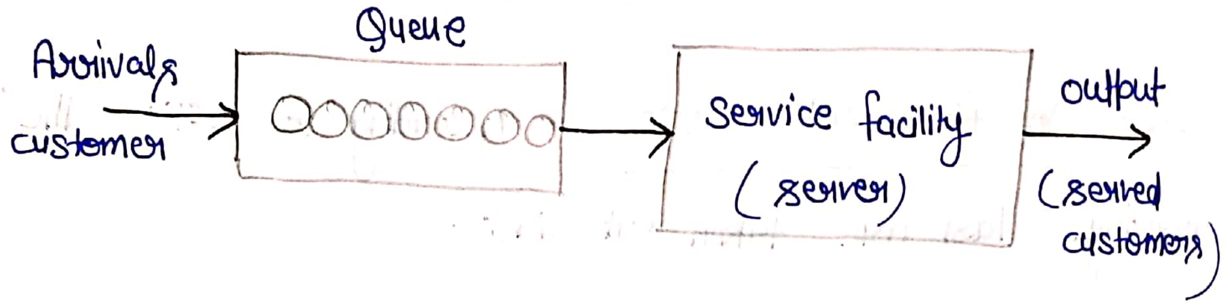
for example —

iv) Priority Based :- under this rule, customers are given priority for service according to particular conditions.

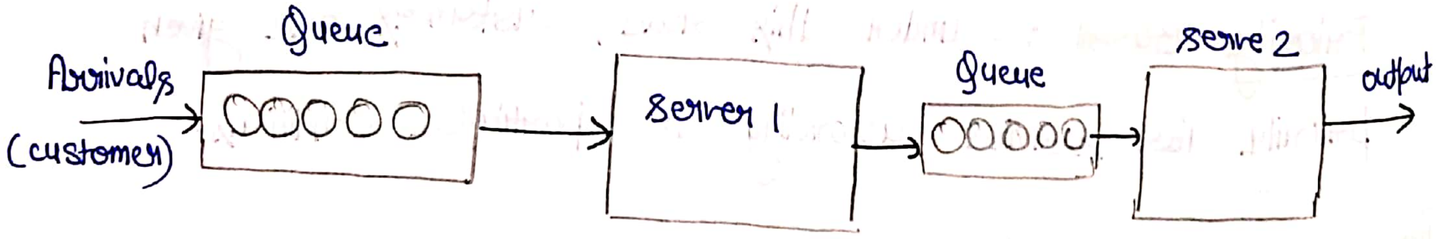
for example — serious patients are given priority for treatment.

4) Service Mechnism :- Service mechnism prescribes the number and configuration of servers. The basic structures of service mechanism are —

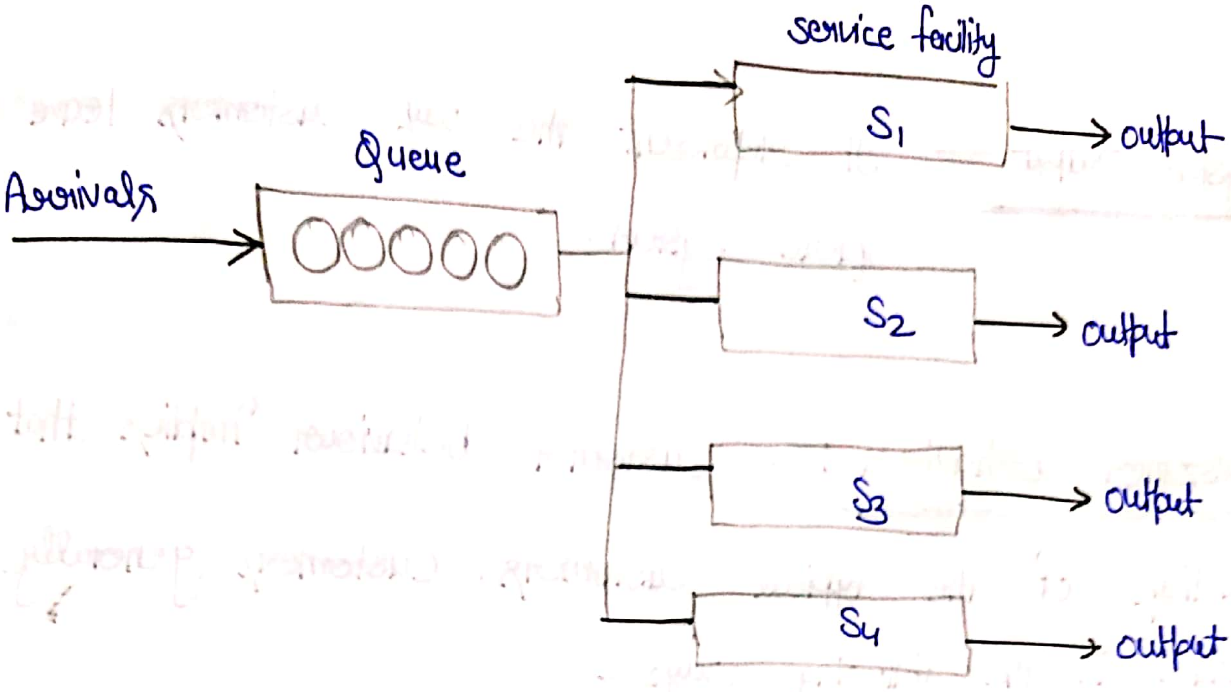
1) Single Queue - Single Server :- In this case the arriving units from one queue to be served by a single service facility.



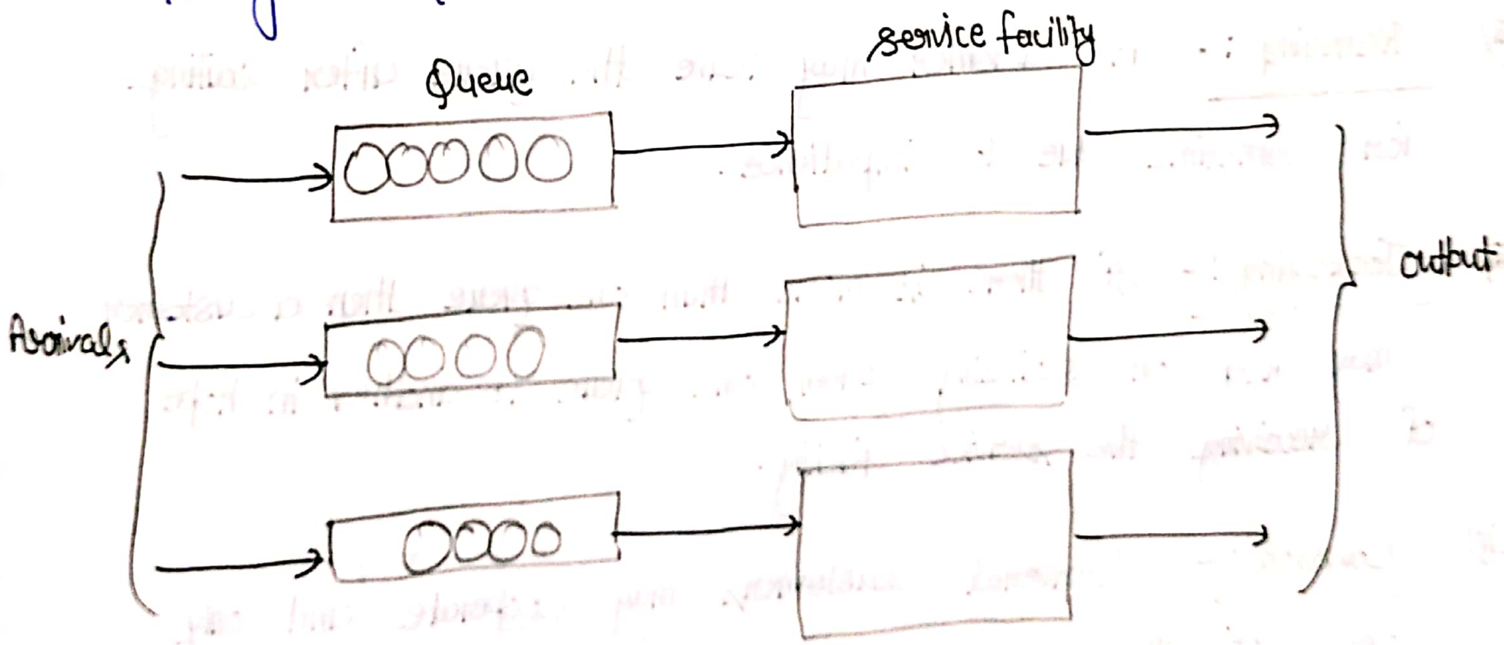
2) Single Queue, Multiple Servers in Series :- In this model, the customers are served at number of servers arranged in series i.e. the customer enters the first station and gets a portion of service and then again moves on nex station ... and so on and finally leaves the system.



3) Single Queue, Multiple Servers in Parallel :- There is more than one server which provides the same type of service facility at parllel manner.



4) <u>Multiple Queue, Multiple server</u> :- In this case, the arrival units from queue and service channel are multiple.



The average number of customers that can be served per unit time is called service rate (μ). Reciprocal of service rate is called service time ($1/\mu$)

5) System Output — It represent the way customers leave the queue system.

Customer Behaviour :— Customer behaviour implies that reactions of the typical customers. Customers generally behave in the following ways—

(i) Balking :— A customer may not like to join the queue due to lack of time or space or long waiting line etc.

(ii) Reneging :— A customer may leave the queue after waiting for sometime due to impatience.

(iii) Jockeying :— If there is more than one queue then a customer may keep on switching from one queue to another in hope of receiving the service facility.

(iv) Collusion :— Serveral customers may cooperate and only one of them may stand in the queue.

Symbols and Notations for Queuing System :-

The notations used in the analysis of a queuing system are as follows —

n = number of customer in the system (waiting and in service)

P_n = probability of n customers in the system

λ = mean arrival rate

μ = mean service rate

$\therefore$ traffic intensity $\rho = \frac{\lambda}{\mu}$ = utilization rate

S = number of service channels

L_q = average number of customer in the queue.

L_s = average number of customer in the system (waiting + being served)

W_q = average waiting time in the queue.

W_s = average waiting time in the queue system (waiting time + service time)

Scanned with CamScanner

State of Queuing System :-

Transient & Steady state :-

The transient state of queuing system is the state where behaviour (like input, output, queuing length etc) are depend upon time.

And a steady state system is the one in which behaviour of the system independent of time.

Let $P_n(t)$ denote the probability that there are n customers in the system at time t (depend upon time)

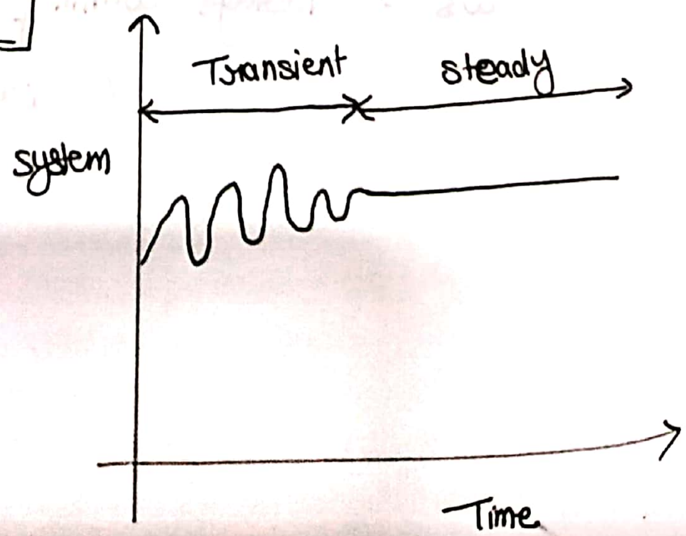
Then in steady state –

$$\lim_{t \to \infty} P_n(t) = P_n \quad (\text{independent of } t)$$

or
$$\lim_{t \to \infty} \frac{d P_n(t)}{dt} = \frac{dP_n}{dt}$$

or
$$\boxed{\lim_{t \to \infty} P_n'(t) = 0}$$

usually, a system in transient state during the early stage of its operation, then go to steady state.



Note – 1) Meaning of Symbolic Representations of Queuing model :-

Symbol of Queuing model is symbolic represented as-

$$(A/B/C) : (p/q/r)$$

where A = Arrival pattern (distribution of intervals between arrivals)

B = Service pattern (distribution of service duration)

C = number of servers or service facility.

p = queuing discipline (like FCFS, SIRO, ... etc)

q = system capacity.

r = population size (number of possible customers)

2) Poisson Distribution – If the Poisson distribution deals with the number of occurrences in a fixed period of time.

The exponential distribution deals with the time between occurrences of successive events as time flows by continuously.

Classification of Queuing Model :- The various type of queuing models are as follows.

1) Probabilistic Queuing Models.
2) Deterministic Queuing Models
3) Mixed Queuing Models.

1) **Probabilistic Queuing Models** :- If each customer arrives at random intervals and service time is known with uncertainty or determined on the basis of probability then it is known as probablistic model. These are following types –

Model I (Erlang Model) :- This model is symbolically represented as– $(M/M/1):(FCFS/\infty/\infty)$

M = Markovian (Poisson) arrival distribution.

M = Markovian (Exponential) service distribution

FCFS = service discipline, first come-first served basis

system capacity (maximum number of customers that allowed) and population size are infinite.

Model II (General Erlang Model) :- (M/M/1) : (FCFS/∞/∞)

It is same as Erlang model, only difference is that in this model arrival and service rate depend upon the length of queue.

(i.e. some persons desiring service but not join the queue because it is too long, that affecting the arrival rate)

Model III - (M/M/1) : (SIRO/∞/∞)

It is same as model I expect that service discipline is SIRO (service in random order).

Model IV - (M/M/1) : (FCFS/N/∞)

In this model capacity of system is limited or finite i.e. only finite number of arrivals.

(New arrivals joins the queue, if the space is available and if queue is full, the arrival leaves the system without joining the queue.)

Model V - (M/M/1) : (FCFS/n/M) :-

It is finite population and limited system capacity model.

Model VI - $(M/M/c) : (FCFS/\infty/\infty)$:-

In this model, multiple service facility is used to provide the ~~serie~~ service which are arranged in parallel manner.

Model VII - $(M/E_K/1) : (FCFS/\infty/\infty)$:-

In this model instead of exponential service time, there is Erlang service time with K phases.

Model VIII $(M/M/1) : (GD/m/n)$:- Here, general service

discipline is used to provide the service. In this model, it is assume that system capacity is less than or equal to the population size i.e. $m \leq n$.

Model IX $(M/M/c)(GD/m/n)$:-

Mode X - This is called power supply model.

2) **Deterministic Queuing Model** :- If each customer arrives at known intervals and the service time is known with certainty.

$(D/D/1) : (FCFS/\infty/\infty)$

3) **Mixed Queuing Model** :- $(M/D/1) : (FCFS/\infty/\infty)$

Model I (Erlang Model Single Channel Queuing Model)

$(M/M/1) : (FCFS/\infty/\infty)$

When one service channel is used to provide the service to customer then it is called single channel queuing model. This is simplest and commonly used queuing model.

Assumptions :-

1) Both the arrivals and service rate are independent of the number of customer in the waiting line.
2) The arrivals occur completely at random according to Poisson distribution.
3) There is only one queue & one service facility.
4) Arrivals are handled on first come- first served (FCFS) basis.
5) The mean service rate is higher than the mean arrival rate (i.e. $\mu > \lambda$)

Important Formula :-

1) Probability that there is no one in the system.

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

2) Probability that the customer will have to wait in queue (i.e. probability that the service facility is busy)

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

3) Probability that there are 'n' customer in the queue system.

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

4) Average expected number of customers in the queue

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

5) Average expected number of customers in the queue system

$$L_s = \frac{\lambda}{\mu-\lambda} = \text{system length}$$

6) Average waiting time in the queue

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

7) Average time in the queue system (waiting time + service time)

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

8) Probability that a customer shall have to wait between t_1 & t_2

$$P = \int_{t_1}^{t_2} (\mu-\lambda) \cdot e^{-(\mu-\lambda)t} \cdot dt$$

Ques - Customers arrive at one person barber shop according to a poisson process with a mean inter-arrival time of 20 minutes. Customers spend on an average of 15 minutes in the barber's chair.

(a) what is the probability that a new arrival need not wait for the barber to be free ?

(b) what is the expected number of customers in the barber shop ?

(c) How much time can a customer expect to wait for his turn ?

(d) How much time can a customer expect to spent in his shop ?

(e) Management will put in another chair and hire another barber when a customer's average time in the shop exceeds 1.25 hours. How much must the average rate of arrivals increase to warrant a second barber ?

Sol - Given that,

$$\text{arrival rate} = \frac{60}{20} = 3/hr = \lambda$$

& $$\text{service rate} = \frac{60}{15} = 4 \text{ per hour} = \mu$$

(a) The probability that a new arrival need not wait for the barber

$$P_0 = 1 - \frac{\lambda}{\mu} = 1 - \frac{3}{4} = \frac{1}{4} = 0.25$$

(b) Expected number of customers in the barber shop (system length)

$$L_s = \frac{\lambda}{\mu - \lambda} = \frac{3}{4-3} = 3$$

(c) The time the customer expects to wait for this turn (average time in the queue)

$$W_q = \frac{\lambda}{\mu(\mu-\lambda)} = \frac{3}{4(4-3)} = \frac{3}{4} \text{ hr} = 0.75 \text{ ~~min~~ hour}$$

$$= 45 \text{ min}$$

(d) The time customer expects to spend in the shop

$$W_s = \frac{1}{\mu - \lambda} = \frac{1}{4-3} = 1 \text{ hr} = 60 \text{ min}$$

(e) Let customers average waiting time = 1.25 hours

then $W_q = 1.25$ i.e.

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

or $$1.25 = \frac{\lambda}{4(4-\lambda)}$$

or $$1.25 = \frac{\lambda}{16 - 4\lambda}$$

or $$20 - 5\lambda = \lambda \Rightarrow 20 = 6\lambda$$

$$\Rightarrow \lambda = \frac{20}{6} = 3.33 \text{ hours}$$

Ques Patrons arrive at a reception counter at an average inter-arrival rate of 2 minutes. The receptionist in duty takes on average of one minute per patron.

(a) what is the chance that a patron will straight way meet the receptionist.

(b) For what portion of time the receptionist is busy?

(c) what is the average queue length.

(d) what is the average number of patrons in the system?

(e) what is the average waiting time of a patron?

(f) what average time a patron spends in the system?

(g) Suppose management wants to keep a second receptionist when the average waiting time of an arrival exceeds 1.5 minutes. Find what should be the average inter-arrival time to justify a second receptionist?

Single Server and Finite Population (M/M/1) : (FCFS/N/M) :-

Let M denote the finite number of customers in the source, λ is customer arrival rate (individual) and μ be the service rate.

Formula –

1) Probability that the system shall be idle :-

$$P(0) = \frac{1}{\sum_{i=0}^{M}\left[\frac{M!}{(M-i)!}\left(\frac{\lambda}{\mu}\right)^{i}\right]+1}$$

2) Probability that there shall be n customers in the system :

$$P_n = P(0)\left(\frac{\lambda}{\mu}\right)^{n}\frac{M!}{(M-n)!} \qquad 0 < n \leq M$$

3) Expected length of the queue

$$L_q = M - \frac{\lambda+\mu}{\lambda}(1-P(0))$$

4) Expected number of customers in the system :

$$L_s = L_q + [1-P(0)]$$

5) Expected waiting time of a customer in the queue

$$W_q = \frac{L_q}{\mu[1-P(0)]}$$

5) Expected time a customer spends in the system

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

Ques – A factory has five machines. On an average, there are two machine breakdowns every 5 weeks. Assuming the repairing capacity is one machine per week, the repairing time being exponentially distributed, determine –

1) The probability that the service facility will be idle,
2) The probability that there shall be exactly 3 machines to be, and being repaired.
3) The expected length of the queue.
4) The expected number of machines waiting to be and being repaired.
5) The expected time a machine shall wait in the queue to be repaired and
6) The expected time that a machine shall spend in the system i.e. waiting in and getting repaired.

Sol – According to the questions –

$M = 5$ machines

$\mu = 1$ machine per week

$\lambda = \frac{2}{5} = 0.4$ machine per week

$\therefore \quad \rho = \frac{\lambda}{\mu} = \frac{0.4}{1} = 0.4 < 1$

(i) The probability that the service facility will be idle

$$P(0) = \frac{1}{\sum_{i=0}^{M}\left[\frac{M!}{(M-i)!}\left(\frac{\lambda}{\mu}\right)^i\right]}$$

calculation of P(0) —

i	$\frac{M!}{(M-i)!}$	$\left(\frac{\lambda}{\mu}\right)^i = (\rho)^i$	$\frac{M!}{(M-i)!}\left(\frac{\lambda}{\mu}\right)^i$
0	1	$(0.4)^0 = 1$	1
1	5	$(0.4)^1 = 0.4$	2
2	20	$(0.4)^2 = 0.16$	3.20
3	60	$(0.4)^3 = 0.064$	3.84
4	120	$(0.4)^4 = 0.0256$	3.072
5	120	$(0.4)^5 = 0.01024$	1.2288
			Total = 14.3408

So,

$$P(0) = \frac{1}{14.3408} = 0.0697$$ Ans

2) The probability that there shall be exactly 3 machines in the system -

$$P_3 = P(0)\left(\frac{\lambda}{\mu}\right)^3 \frac{M!}{(M-3)!}$$

Or $P_3 = 0.0697\,(0.4)^3 \times \frac{5!}{(5-3)!}$

$= 0.0697 \times 0.064 \times 60 = 0.2676$ Ans

3) The expected length of the queue

$$L_q = M - \frac{\lambda+\mu}{\lambda}\,[1-P(0)]$$

$$= 5 - \frac{0.4+1}{0.4}\,[1-0.0697]$$

$$= 5 - 3.2560 = 1.7439 \text{ machines}$$ Ans

4) The expected number of machines in the system

$$L_s = L_q + [1-P(0)]$$

$$= 1.7439 + (1-0.0697) = \cancel{2.6742}\ 2.6742 \text{ machines}$$ Ans

5) The expected time a machine has to wait in the queue

$$W_q = \frac{L_q}{\mu\,[1-P(0)]} = \frac{1.7439}{1\,[1-0.0697]} = 1.8745 \text{ weeks}$$ Ans

6) The expected time a machine spends in the system

$$W_s = W_q + \frac{1}{\mu} = 1.8745 + \frac{1}{1} = 2.8745 \text{ weeks.}$$ Ans

Scanned with CamScanner

Ques - Past records indicate that of the five machines that a factory owns, breakdown occurs at random and the average time between the breakdown is 2 days. Assuming that the repairing capacity of the workman is one machine a day and the repairing is distributed exponentially, determine the following -

1) The probability that the service facility will be idle.
2) The probability of various numbers of machine (0 through 5) to be and being repaired.
3) The expected length of the queue.
4) The expected number of machines waiting to be, and being repaired.
5) The expected time that a machine shall wait in the queue to be repaired.
6) The expected time that a machine will be in the system.

Sol Here $M = 5$ machines

$\mu = 1$ machine / day

$\lambda = \frac{1}{2} = 0.5$ machine / day

$\therefore \quad \rho = \frac{\lambda}{\mu} = 0.5$

Multi Channel Queuing Model :- $(M/M/C):(FCFS/\infty/\infty)$

In this model, multiple service channel provide the service and arranged in parallel order.

Assumptions –

1) There are ~~server~~ several servers in the service facility. The servers are arranged in parallel.
2) The customers wait in the single queue until any one of the server becomes free to provide service.
3) The service discipline in FCFS i.e. first come first served basis.
4) The arrival rate λ and the service rate μ are mean values from Poisson distribution and exponential distribution respectively.
5) The service rate is same for each server.
6) The queue is of infinite length.

Important Formulae :-

1) Utilization rate

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

2) Probability that there is no customer in the system

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

3) Probability that there are n customers in the system (P_n)

$$P_n = \left(\frac{\lambda}{\mu}\right)^n \cdot P_0 \quad \text{for } n < c$$

4) Average number of customers in the queue

$$L_q = \frac{\lambda \cdot \mu \left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\,(c \cdot \mu - \lambda)^2} \cdot P_0$$

5) Average number of customers in the system

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

or

$$L_s = \frac{\lambda \cdot \mu \left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\,(c \cdot \mu - \lambda)^2} P_0 + \frac{\lambda}{\mu}$$

6) Average waiting time in the queue

$$W_q = \frac{\mu \left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\,(c\mu - \lambda)^2} P_0$$

i.e. $W_q = \frac{L_q}{\lambda}$

7) Average waiting time in the system

$$W_s = \frac{\mu \cdot \left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\,(c\mu-\lambda)^2} P_0 + \frac{1}{\mu}$$

i.e. $$W_s = \frac{L_s}{\lambda}$$

8) Probability that a customer has to wait

$$P(n \geq c) = \frac{\mu \left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\,(c\mu-\lambda)} \cdot P_0$$

Example - A post-office has 3 windows providing the same service. It receives on an average 30 customers per hour. Arrivals are Poisson distributed and service time is exponentially distributed. Each window serves as an average 12 customers per hour.

1) what is the probability that a customer will be served immediately?
2) what is the probability that a customer will have to wait?
3) what is the average number of customers in the system?
4) what is the average total time a customer must spend in the post office?

Sol - Now, $c = 3$, $\lambda = 30$ customers per hour

$\mu = 12$ customers per hour

and $\frac{\lambda}{\mu} = \frac{30}{12} = \frac{5}{2}$

(1) The probability that a customer will be served immediately -

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

or

$$P_0 = \left[\left(1+\frac{\lambda}{\mu}+\frac{1}{2}\left(\frac{\lambda}{\mu}\right)^2\right)+\left(\frac{1}{6}\left(\frac{\lambda}{\mu}\right)^3 \cdot \frac{3\mu}{3\mu-\lambda}\right)\right]^{-1}$$

$$= \left[\left(1+\frac{5}{2}+\frac{25}{8}\right)+\left(\frac{1}{6}\times\frac{125}{8}\times\frac{3\times 12}{3\times 12-30}\right)\right]^{-1}$$

$$= \left[\frac{53}{8}+\frac{125}{8}\right]^{-1} = \left(\frac{178}{8}\right)^{-1}$$

$$= \frac{8}{178} = 0.0449$$

(2) The probability that a customer will have to wait

$$P(n \geq c) = \frac{\mu\left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\,(c\mu-\lambda)} \cdot P_0$$

$$= \frac{12\left(\frac{5}{2}\right)^3}{(3-1)!\,(3\times 12-30)} \cdot P_0$$

$$P(n \geq c) = \frac{12 \times \frac{125}{8}}{2 \times 6} \times \frac{8}{178}$$

$$= \frac{125}{178} = 0.702$$

3) Average number of customers in the system

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

Now,

$$L_q = \frac{\lambda \cdot \mu \left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\,(c\cdot\mu-\lambda)^2} \cdot P_0$$

$$= \frac{30 \times 12 \times \left(\frac{30}{12}\right)^3}{(3-1)!\,(3 \times 12 - 30)^2} \cdot \frac{8}{178}$$

$$= \frac{360 \times \left(\frac{5}{2}\right)^3}{2!\,(6)^2} \cdot \frac{8}{178}$$

$$= \frac{360 \times \frac{125}{8}}{2 \times 36} \cdot \frac{8}{178} = \frac{625}{178}$$

So,

$$L_s = \frac{625}{178} + \frac{5}{2} = 3.51 + 2.5$$

$$= 6.01 \approx 6 \text{ customer}$$

4) The average total time a customer spends in the post office –

$$W_s = \frac{\mu\left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\,(c\mu-\lambda)^2} P_0 + \frac{1}{\mu}$$

$$= \frac{12\left(\frac{5}{2}\right)^3}{(3-1)!\,(3\times12-30)^2} \cdot \frac{8}{178} + \frac{1}{12}$$

$= 0.20$ hrs.

or

$$W_s = \frac{L_s}{\lambda}$$

$$= \frac{6}{30} = \frac{1}{5} = 0.20 \text{ hrs}$$

$= 12$ mins Ans

Ques – Sunil Medicals is manned by three salesman. Any saleman can provide desired service to any customer. The customers arrive at the counter according to Poisson distribution at an average rate of 30 per hour. The service time is exponential with a mean rate of 3 minutes?

(a) what fraction of time are all three attendants busy?

(b) what is the mean number of customers waiting to be attended?

(c) what average time does a customer spend at the shop?

(d) what is the probability that a customer has to wait?