18MAB204T
PQT Unit-4

Queuing Theory

Queuing theory is the mathematical study of queuing, or waiting in lines. Queues contain customers (or “items”) such as people, objects, or information. Queues form when there are limited resources for providing a service. For example, if there are 5 cash registers in a grocery store, queues will form if more than 5 customers wish to pay for their items at the same time.

A basic queuing system consists of an arrival process (how customers arrive at the queue, how many customers are present in total), the queue itself, the service process for attending to those customers, and departures from the system.

Mathematical queuing models are often used in software and business to determine the best way of using limited resources. Queueing models can answer questions such as: What is the probability that a customer will wait 10 minutes in line? What is the average waiting time per customer?

The following situations are examples of how queueing theory can be applied:

- Waiting in line at a bank or a store
- Waiting for a customer service representative to answer a call after the call has been placed on hold
- Waiting for a train to come
- Waiting for a computer to perform a task or respond
- Waiting for an automated car wash to clean a line of cars

Characterizing a Queuing System

Queuing models analyze how customers (including people, objects, and information) receive a service. A queuing system contains:

- **Arrival process.** The arrival process is simply how customers arrive. They may come into a queue alone or in groups, and they may arrive at certain intervals or randomly.
- **Behavior.** How do customers behave when they are in line? Some might be willing to wait for their place in the queue; others may become impatient and leave. Yet others might decide to rejoin the queue later, such as when they are put on hold with customer service and decide to call back in hopes of receiving faster service.
- **How customers are serviced.** This includes the length of time a customer is serviced, the number of servers available to help the customers, whether customers are served one by one or in batches, and the order in which customers are serviced, also called **service discipline.**
- **Service discipline**

  FCFS(First come first serve), FIFO(First In First Out), FILO(First In Last Out), SIRO(Service In Random Order) and SIP(Service In Priority)
- **Waiting room.** The number of customers allowed to wait in the queue may be limited based on the space available.

Mathematics of Queuing Theory

Kendall's notation is a shorthand notation that specifies the parameters of a basic queuing model. Kendall's notation is written in the form (a/b/c):(d/e), where each of the letters stand for different parameters.

Where

- 'a' represents arrival pattern of the customer.
- 'b' represents service pattern of the customer.
- 'c' represents number of servers.
- 'd' represents capacity of the system.
- 'e' represents queue discipline.

Little's Formula,

which was first proven by mathematician John Little, states that the average number of items in a queue can be calculated by multiplying the average rate at which the items arrive in the system by the average amount of time they spend in it.In mathematical notation, the Little's law is: $L = \lambda W$.

Where L is the average number of items,
λ is the average arrival rate of the items in the queuing system,
W is the average amount of time the items spend in the queuing system.

(i) $L_s = \lambda W_s$ (ii) $L_q = \lambda W_q$ (iii) $W_s = W_q + \frac{1}{\lambda}$ (iv) $L_s = L_q + \frac{\lambda}{\mu}$

Model I (M/M/1) : (∞ / FIFO) (Single server , infinite capacity queue)

This is a simple queue with poisson arrival, exponential service time , single server , infinite capacity and First in First out queue discipline.

The arrival rate and service rate are constant (λ and μ res.) To find the steady state probabilities, put $\lambda_{n=} \lambda$, $\mu_n = \mu$ for

all n in $P_0 = \frac{1}{1+\sum_{n=1}^{\infty}\frac{\lambda_0\lambda_1\lambda_2 \ldots \lambda_{n-2}\lambda_{n-1}}{\mu_1\mu_2\mu_3 \ldots \mu_{n-1}\mu_n}}$ and $P_n = \frac{\lambda_0\lambda_1\lambda_2 \ldots \lambda_{n-2}\lambda_{n-1}}{\mu_1\mu_2\mu_3 \ldots \mu_{n-1}\mu_n} P_0$

=> $P_0 = 1 - \frac{\lambda}{\mu}$ and $P_n = \left(\frac{\lambda}{\mu}\right)^n P_0$ => $P_n = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right)$

Let , the average arrival rate of customer = λ and the average service rate of the server = μ and $\lambda < \mu$.

P(server is busy) = Traffic intensity = $\rho = \frac{\lambda}{\mu}$ = P(system is busy)

P (server is idle) = P(No customer in the system) = P(system is empty) =

$$P_0 = 1 - \rho \quad => P_0 = 1 - \frac{\lambda}{\mu}$$

Steady state Probabilities :

$$P_0 = 1 - \frac{\lambda}{\mu} \quad => P_0 = 1 - \rho \quad \text{and } P_n = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right) \quad => P_n = \rho^n (1 - \rho)$$

Characteristics of the queue :

1.Steady state probability p_0

$$p_0 = 1 - \frac{\lambda}{\mu}$$

$$p_n = \left(\frac{\lambda}{\mu}\right)^n p_0 = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right)$$

2.Average no. of customers in the system L_s

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

3.Average number of customers in the queue (L_q)

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

4. Average waiting time of a customer in the system W_s

$$W_s = \frac{L_s}{\lambda} = \frac{\lambda}{\lambda(\mu - \lambda)} = \frac{1}{(\mu - \lambda)}$$

5. Average waiting time of a customer in the queue Wq

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

Characteristics of the queue :

6. **Probability that the no. of customers in the system exceeds 'k'**

$$p(N > k) = \left(\frac{\lambda}{\mu}\right)^{k+1}$$

7. **Average number L_w of customers in the non-empty queue (L_w)**

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

8. **Average waiting time of customers in the queue, if he has to wait.**

$$E\left(W_q/W_q > 0\right) = \frac{1}{\mu-\lambda}$$

9. **Probability density function (pdf) of the waiting time in the system**

$$f(w) = (\mu - \lambda)e^{-(\mu-\lambda)w}$$

10. **Probability that the waiting time of a customer in the system exceeds 't'**

$$g(w) = \begin{cases} \frac{\lambda}{\mu}(\mu - \lambda)e^{-(\mu-\lambda)w}, w > 0 \\ 1 - \frac{\lambda}{\mu}, w = 0 \end{cases}$$

Problem 1: Customers arrive at a watch repair shop according to a poisson process at a rate of one per every 10 minutes and the service time is an exponential r.v.with mean 8 min.

(1) Find the average number of customers Ls in the shop
(2) Find the average time a customer spends in the shop Ws
(3) Find the average number of customers in the queue Lq
(4) What is the probability that the server is idle.

Solution: Watch repair shop.

Nothing given about the no. of machines.

∴assume only one machine (1 server)

No restriction about the accommodation of customers

$\therefore (M/M/1):(\infty/FIFO)$ model.

For 10 min. 1 customer arrived.

$\therefore \, for 1\, min = \frac{1}{10}\, customer\ arrive \Rightarrow \lambda = \frac{1}{10}\, per\ min.$

8 min. service time for 1 customer.

$\therefore \, for 1\ min = \frac{1}{8}\, customer \Rightarrow \mu = \frac{1}{8}\, per\ min.$

(i) $L_s = \frac{\lambda}{\mu-\lambda} = \frac{\frac{1}{10}}{\frac{1}{8}-\frac{1}{10}} = 4$ Customers

(ii) $W_s = \frac{1}{\mu-\lambda} = \frac{1}{\frac{1}{8}-\frac{1}{10}} = 40$ minutes

(iii) $L_q = \frac{\lambda^2}{\mu(\mu-\lambda)} = \frac{\left(\frac{1}{10}\right)^2}{\frac{1}{8}\left(\frac{1}{8}-\frac{1}{10}\right)} = \frac{16}{5}$ = 3.2 ≈ 3 customers

(iv) $p_0 = server\ is\ idle\ when\ no\ customer\ is\ in\ the system$

$= 1 - \frac{\lambda}{\mu} = \frac{1}{5}$ = 0.2

Problem-2 : Customers arrive at a one – man barber shop according to a poisson process with a mean inter arrival time of 20 min. Customers spend an average of 15 min. in the barber chair. If an hour is used as a unit time , then

(i) What is the probability that a customer need not wait for a hair cut ?

(ii) What is the expected no. of customers in the barber shop and in the queue?

(iii) How much time can a customer expect to spend in the barber shop?

(iv) Find the average time that the customer spend in the queue.

(v) What is the probability that there will be 6 or more customers waiting for service?

Solution: inter arrival time 20 min $\Rightarrow for 20\ min = 1 customer$

$$\therefore for 1\ min = \frac{1}{20} customer \Rightarrow \lambda = \frac{1}{20} per min.$$

15 min. service time for 1 customer.

$$\therefore for\ 1\ min = \frac{1}{15} customer \Rightarrow \mu = \frac{1}{15} per\ min.$$

(i) If no customer in the system, the customer no need to wait

$$\therefore p(n=0) = 1 - \frac{\lambda}{\mu} = 1 - \frac{\frac{1}{20}}{\frac{1}{15}} = \frac{1}{4}$$

(ii) Barber shop + queue $\Rightarrow$ $system$

$$\therefore E(N_s) = L_s = \frac{\lambda}{\mu - \lambda} = \frac{\frac{1}{20}}{\frac{1}{15} - \frac{1}{20}} = 3$$

(iii) Time to spend in the barber shop (system)

$$\therefore W_s = \frac{1}{\mu - \lambda} = \frac{1}{\frac{1}{15} - \frac{1}{20}} = 60 min.$$

(iv) $$W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{\frac{1}{20}}{\frac{1}{15}\left(\frac{1}{15} - \frac{1}{20}\right)} = 45 min.$$

(v) $$p(n \geq 6) = \left(\frac{\lambda}{\mu}\right)^6 = \left(\frac{1/20}{1/15}\right)^6 = 0.1779$$