ROHINI COLLEGE OF ENGINEERING & TECHNOLOGY

DEPARTMENT OF MATHEMATICS

LECTURE NOTES ON

BA4201 / QUANTITATIVE TECHNIQUES FOR DECISION MAKING

UNIT V : QUEUING THEORY

INTRODUCTION TO QUEUING THEORY

INTRODUCTION

Queueing or waiting in a line is a common situation occurring in everyday life. We wait in queues in ticketbooths, bus stops, post offices, banks, traffic lights and so on. In general, a queue is formed when there are customers who require some sort of services and the queuing problem is identified by the presence of a group of customers who arrive randomly to receive some service. The customer may be a person, machine, vehicle or anything else which requires service. The objective of a queuing model is to find out the optimum service rate and the number of servers so that the average cost of being in queuing system and the cost of service are minimized. The objective of a queueing system is to find ways of reducing the time spent in waiting by the customerand at the same time optimizing the cost to the service provider.

BASIC CONCEPTS OF QUEUEING THEORY

Sample point: An outcome of an experiment is called a sample point.

Sample Space: The set of all possible outcomes of a random experiment is called sample space.

Events: Subsets of the sample space are called events.

Random Variable: A random variable is a function that associates a point of the sample space with a realnumber.

Random Process/Stochastic Process: A random process or stochastic process is a family (or collection)of random variables.

FUNDAMENTAL STRUCTURE OF A QUEUING SYSTEM

Queuing theory is related with the mathematical study of queues or waiting lines, a queue is formed whenthere are customers who require some sort of services and the current demand for a service exceeds the current capacity to provide the service. Generally, the customer's arrival and their service time are not known in advance or can't be predicted accurately. Since arrival-departure process are random. So, queuing models developed to reduce waiting time/excessive costs and work for maintaining balance between service capacity and waiting time.

A simple queuing system can be described as follows:

1. Input or arrival process of customers
2. Service mechanism (or process)
3. Queue discipline

The fundamental structure of a queueing system shown in the figure given below

Now we explain all the components of queueing system.

Input or Arrival Process of Customers

The rate at which the customers arrive at the service facility is determined by the arrival process. An input/arrival process can be defined completely by its size, the arrival time distribution, and the attitude of the customers. We describe these, in brief.

i) **Size:** It may be finite or infinite according as the arrival rate is affected or not affected by thenumber of customers in the service system.

ii) **The arrival time distribution:** Mostly, the arrival time distribution is approximated by Poissondistribution.

iii) Customer or arrivals behavior:

- A customer who stay in the system until served no matter how much he has to wait for service. Such a customer is called Patient Customer**.**
- The customer who wait for a certain time in the queue and leave the system without getting service. This kind of customer is known as impatient or reneging behavior.

- If a customer before joining the system get discouraged by seeing the number of customers already in the queue is too large and does not join the queue. This behavior ofthe customer is called Balking behavior.
- Customers who move from one queue to another because they think that their queue ismoving slower, the behavior of the customer is known as queue jockeying.

Service Mechanism (or Process)

Service time distributions are generally exponential distributions. It may be any one of the following types:

1. Single channel facility
2. One queue-several station facilities
3. Several queues-one service station
4. Multi-channel facility and
5. Multistage channel facility

Queue Discipline

Queue discipline is the order or the manner in which the service station selects the next customer from thewaiting line to be served. There are many ways in which a customer to be selected for service. Some of these are described as follows:

i) FIRST IN, FIRST OUT (FIFO) or in other words First Come First Served (FCFS);

ii) Last in First Out (LIFO); and

iii) Service in Random Order (SIRO)

OPERATING CHARACTERISTICS OF A QUEUEING SYSTEM

There are two types of solutions of these equations: (i) **Transient** (ii) **Steady state**.

Transient solutions: The time dependent solutions are known as transient solutions.

Steady state solutions: These solutions are independent of time and represent the probability of the system being in a particular state in the long run.

6.1.1. Operating characteristics/ Performance Measures

Performance Measures of a queuing system are determined by two statistical properties, namely, theprobability distribution of inter-arrival times and the probability distribution of service times. Some of the operating characteristics/performance measures of any queueing system that are of generalinterest for analyzing the system are listed below:

$\mathbf{L_s}$: The average number of customers in the queuing system (those waiting to be

served and those beingserved).

W_s:The average time each customer spends in the queuing system from entry into the queue tocompletion of the service (the time spent waiting in the queue and during the service).

L_q: The average number of customers in the queue waiting to get service (this excludes customersundergoing service).

W_q: The average time each customer spends in the queue waiting to get service (this excludes customerstime spent during the service).

Service idle time: The relative frequency with which the service system is idle.

Classification of Queueing Models

Generally, queueing models are described by five symbols such as a/b/c: d/e or a/b/c/d/e. The first symbol'a' describes the arrival process. The second symbol 'b' describes the service time distribution. The third

symbols 'c' stands for the number of servers. The symbols 'd' and 'e' stand for the system capacity and queue discipline respectively.

First three symbols i.e. a/b/c in the above notation were described by D. Kendall in 1953. Later, A. Lee in1966 added the fourth (d) and fifth (e) to the Kendall notation.