CHAPTER 1

OPTIMIZATION MODELS

Unit-A

UNIT STRUCTURE

1A.1 Introduction to Optimization Models

1A.2 Definitions of Optimization Models

1A.3 Features of Optimization Models

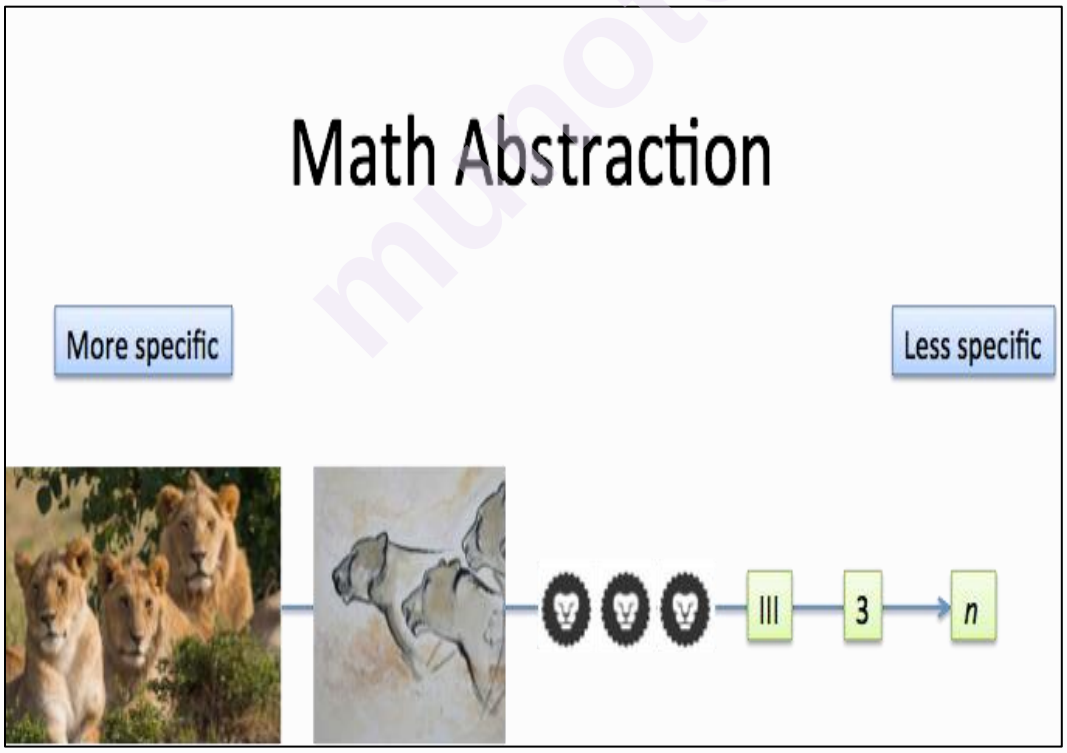
1A.4 Approaches to Optimization Models

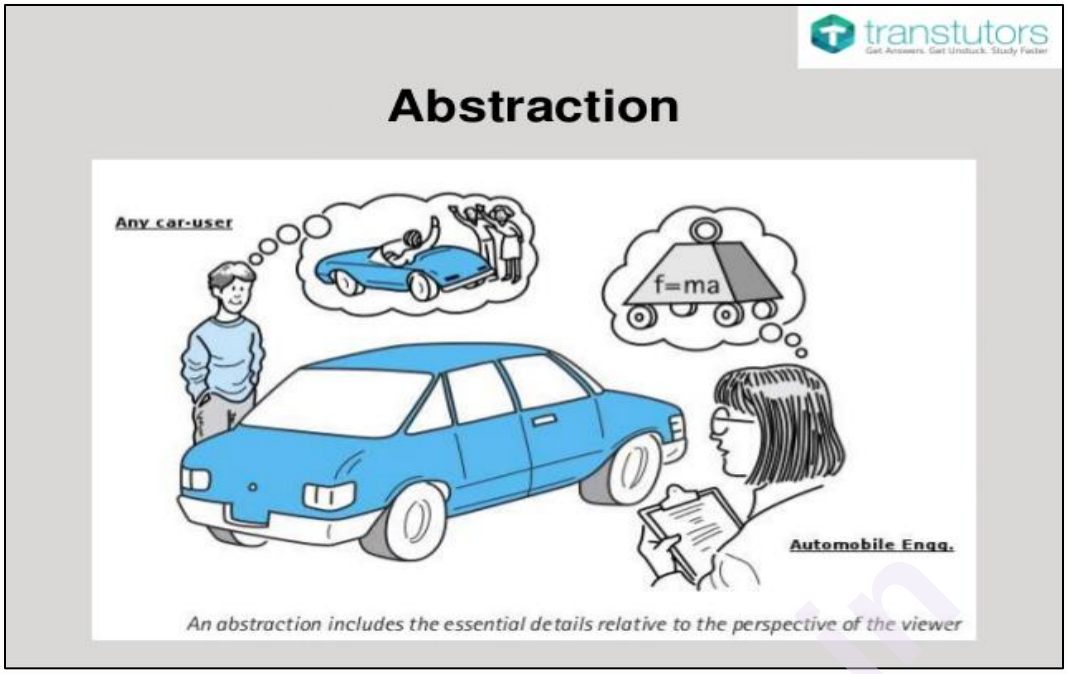
1A.5 Exercise

munotes.in

1A.1. INTRODUCTION TO OPTIMIZATION MODELS

Modeling is the essence of operation research. A model is an abstraction of idealized representation of a real life problem. Modeling is a real life situation helps us to study the different behavior of the problem corresponding to the description of the problem. A model can be a picture, a map, a curve or an equation. The reliability of the decision drawn from the model may depend upon the validity of the model or the basic assumptions on which the model is built. Modeling a real life situation helps us to study the different behavior of the problem corresponding to the description of the problem. Great efforts have been taken by expert to model business situations, military operations, motions of planets and stars and so on. A model is an abstraction or an idealized representation of a real life problem. The objective of a model is to provide means for analyzing the behaviors of the system for further improvement. A map of multiple activity chart, a project network, the representation of the behaviors of a queuing system, a model to forecast the future and the present factors of time series, etc. are all examples a based on the past etc. are all examples of models.





An iconic model is a physical representation of a real project on a different scale. A toy automobile is an iconic model of a real automobile. A globe is an iconic model of the earth. A photograph is also an iconic model. In general, they enlarge or reduce the size of a real system. In other words, they are the images of a real object.

Optimization is central to any problem involving decision making, whether in engineering or in economics or in management. The task of decision making entails choosing between various alternatives. This choice is governed by our desire to make the "best" decision. The measure of goodness of the alternatives is described by an objective function or performance index. Optimization theory and methods deal with selecting the best alternative in the sense of the given objective function.

The area of optimization has received enormous attention in recent years, primarily because of the rapid progress in computer technology, including the development and availability of user-friendly software, high-speed and parallel processors, and artificial neural networks.

An iconic model is a physical representation of a real project on a different scale. A toy automobile is an iconic model of a real automobile. A globe is an iconic model of the earth. A photograph is

also an iconic model. In general, they enlarge or reduce the size of a real system. In other words, they are the images of a real object.

Advantages of Model:

- An iconic model is concrete.
- It is easy to construct the model.
- It is easy to study the model than the system itself.

Disadvantages of Model:

- This model is not suited for further manipulation.
- It cannot be used to study the changes in the operation of the system.
- It is not possible to make any modification of the model.
- Adjustment with changing situations cannot be done in this model.

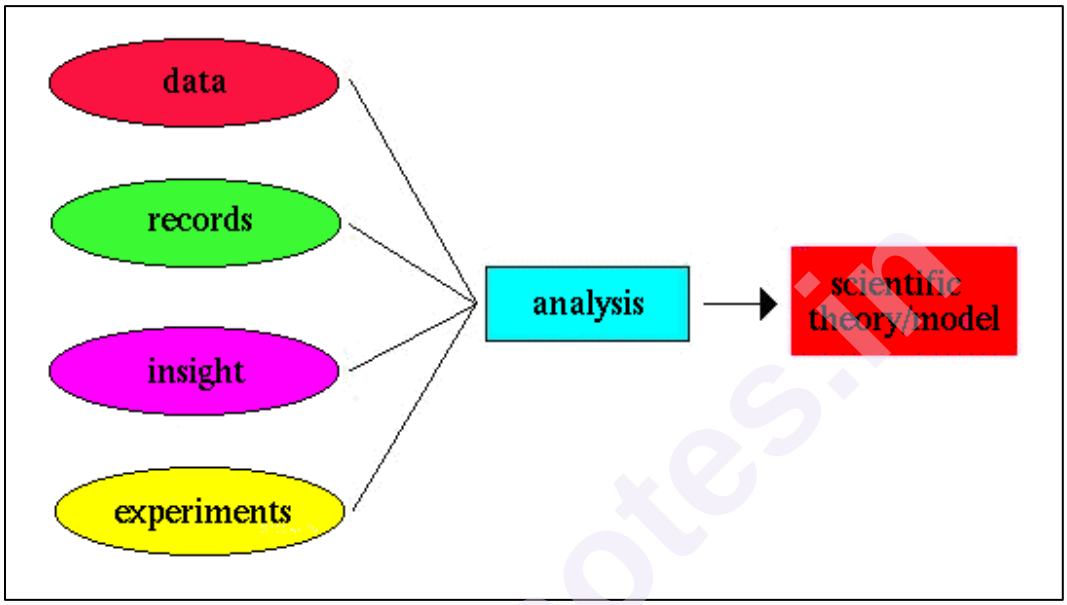
A model is a vehicle for arriving; a well-structured problem of reality. A commentator of a cricket match describes the play as model to enable us to predict the future course of events of the play. It is a descriptive model available for further analysis. It is not always possible to analyze a situation only with the description of the situation. We have to formulate the problem into concrete mathematical representation in the form of a curve, graph or equations. Models could be classified as iconic model, analogue model and symbolic model.

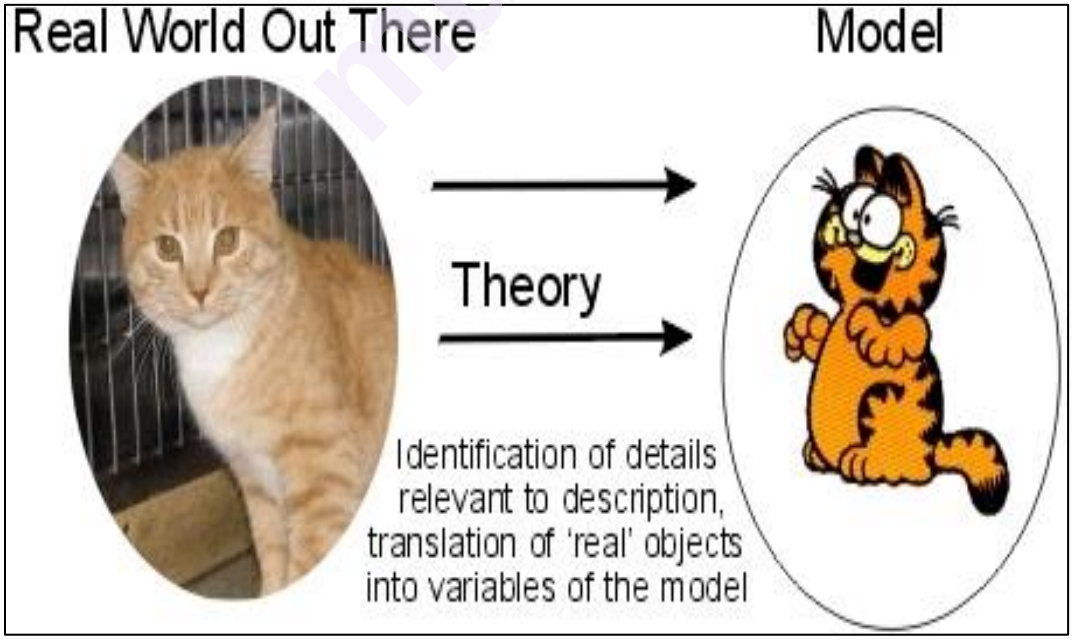
munotes.in

1A.2. DEFINITION OF OPTIMIZATION MODELS

Optimization is the act of obtaining the best result under given circumstances. In design, construction, and maintenance of any engineering system, engineers and business managers have to take many technological and managerial decisions at several stages. The ultimate goal of all such decisions is either to minimize the effort required or to maximize the desired benefit. Since the effort required or the benefit desired in any practical situation can be expressed as a function of certain decision variables, ***optimization can be defined as the process of finding the conditions that give the maximum or minimum value of a function***.

One possible definition of optimization models is "Mathematical Models designed to help institutions and individuals decide *how to allocate scarce resources*, *to activities and to make best use of their circumstances.* Hence optimization models are mathematical models designed to help the managers make better decisions.



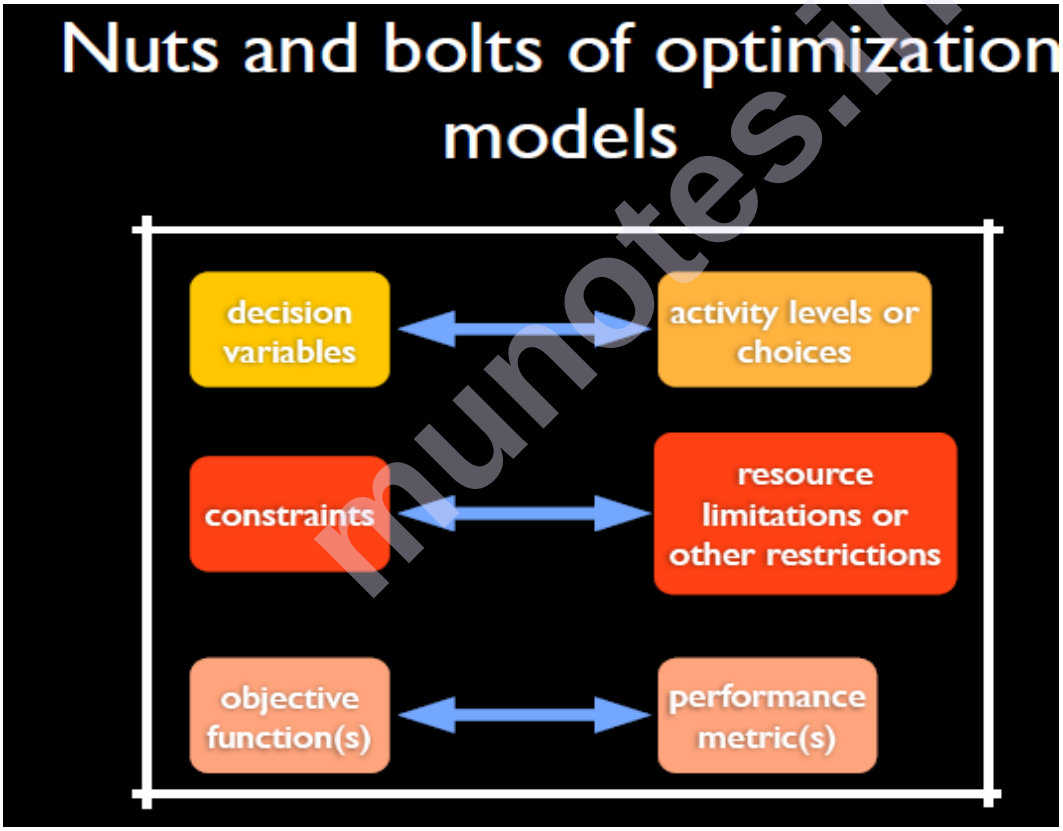


1A.3. FEATURES OF OPTIMIZATION MODELS

An optimization model has three main components:

- An objective function. This is the function that needs to be optimized.
- A collection of decision variables. The solution to the optimization problem is the set of values of the decision variables for which the objective function reaches its optimal value.
- A collection of constraints that restrict the values of the decision variables.

The Optimization Model class provides a common API for defining and accessing variables and constraints, as well as other properties of each model.



munotes.in

The existence of optimization methods can be traced to the days of Newton, Lagrange, and Cauchy. The development of differential calculus methods of optimization was possible because of the contributions of Newton and Leibnitz to calculus. The foundations of calculus of variations, which deals with the minimization of functional, were laid by Bernoulli, Euler, Lagrange, and Weirstrass.

The method of optimization for constrained problems, which involves the addition of unknown multipliers, became known by the name of its inventor, Lagrange. Cauchy made the first application of the steepest descent method to solve unconstrained minimization problems. Despite these early contributions, very little progress was made until the middle of the twentieth century, when high-speed digital computers made implementation of the optimization procedures possible and stimulated further research on new methods. Spectacular advances followed, producing a massive literature on optimization techniques. This advancement also resulted in the emergence of several well-defined new areas in optimization theory.

It is interesting to note that the major developments in the area of numerical methods of unconstrained optimization have been made in the United Kingdom only in the 1960s. The development of the simplex method by Dantzig in 1947 for linear programming problems and the annunciation of the principle of optimality in 1957 by Bellman for dynamic programming problems paved the way for development of the methods of constrained optimization. Work by Kuhn and Tucker in 1951 on the necessary and sufficiency conditions for the optimal solution of programming problems laid the foundations for a great deal of later research in nonlinear programming. The contributions of Zoutendijk and Rosen to nonlinear programming during the early 1960s have been significant. Although no single technique has been found to be universally applicable for nonlinear programming problems, work of Carroll and Fiacco and McCormick allowed many difficult problems to be solved by using the well-known techniques of

unconstrained optimization. Geometric programming was developed in the 1960s by Duffin, Zener, and Peterson. Gomory did pioneering work in integer programming, one of the most exciting and rapidly developing areas of optimization. The reason for this is that most real-world applications fall under this category of problems. Dantzig and Charnes and Cooper developed stochastic programming techniques and solved problems by assuming design parameters to be independent and normally distributed.

There is no single method available for solving all optimization problems efficiently. Hence a number of optimization methods have been developed for solving different types of optimization problems. The optimum seeking methods are also known as *mathematical programming*

munotes.in

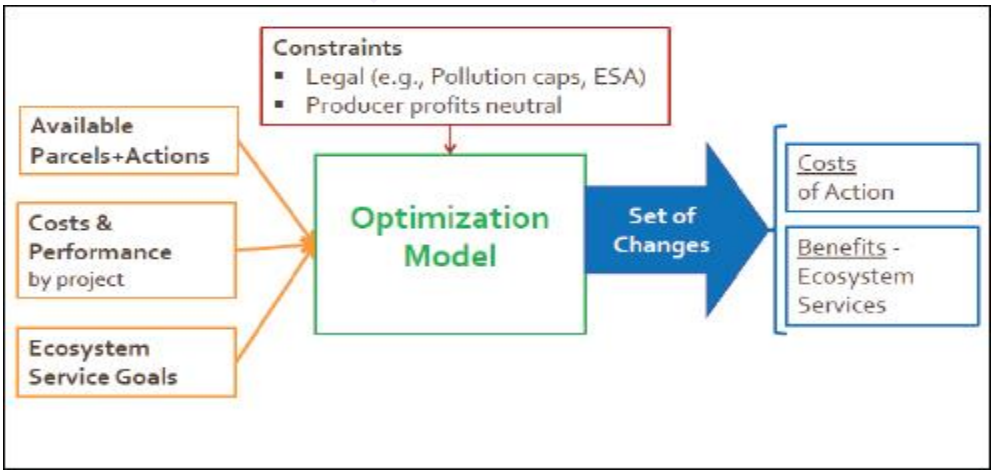
techniques and are generally studied as a part of operations research. *Operations research* is a branch of mathematics concerned with the application of scientific methods and techniques to decision making problems and with establishing the best or optimal solutions. The beginnings of the subject of operations research can be traced to the early period of World War II. During the war, the British military faced the problem of allocating very scarce and limited resources (such as fighter airplanes, radars, and submarines) to several activities (deployment to numerous targets and destinations). Because there were no systematic methods available to solve resource allocation problems, the military called upon a team of mathematicians
to develop methods for solving the problem in a scientific manner. The methods developed by the team were instrumental in the winning of the Air Battle by Britain. These methods, such as linear programming, which were developed as a result of research on (military) operations, subsequently became known as the methods of operations research.

1A.4. APPROACHES TO OPTIMIZATION MODELS

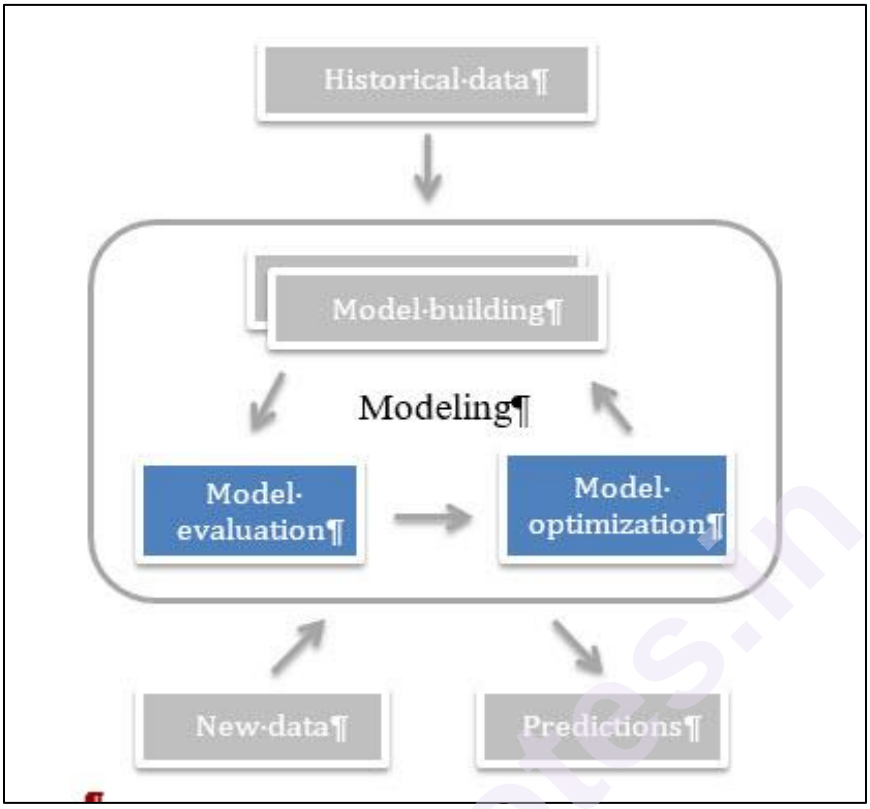
- **Components of an Optimization Model**

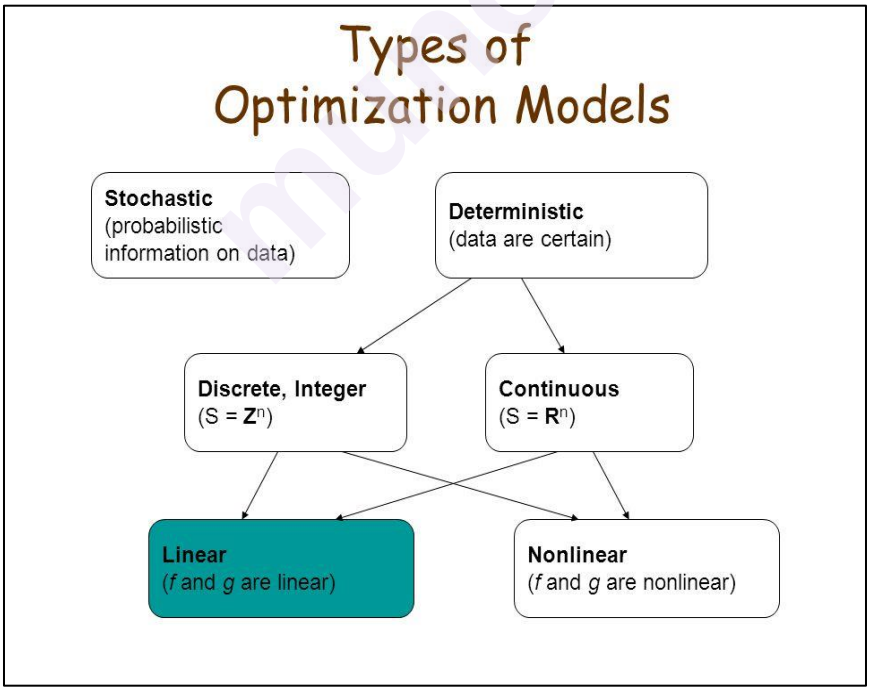
An optimization model has three main components:

- An objective function. This is the function that needs to be optimized.
- A collection of decision variables. The solution to the optimization problem is the set of values of the decision variables for which the objective function reaches its optimal value.
- A collection of constraints that restrict the values of the decision variables.



- **Types of Optimization Models**





Optimization problems can be classified in terms of the nature of the objective function and the nature of the constraints. Special forms of the objective function and the constraints give rise to specialized algorithms that are more efficient. From this point of view, there are four types of optimization problems, of increasing complexity.

An *unconstrained optimization problem* is an optimization problem where the objective function can be of any kind (linear or nonlinear) and there are no constraints. These types of problems are handled by the classes discussed in the earlier sections.

A *linear program* is an optimization problem with an objective function that is linear in the variables, and all constraints are also linear. Linear programs are implemented by the Linear Program class.

A *quadratic program* is an optimization problem with an objective function that is quadratic in the variables (i.e. it may contain squares and cross products of the decision variables), and all constraints are linear. A quadratic program with no squares or cross products in the objective function is a linear program. Quadratic programs are implemented by the Quadratic Program class.

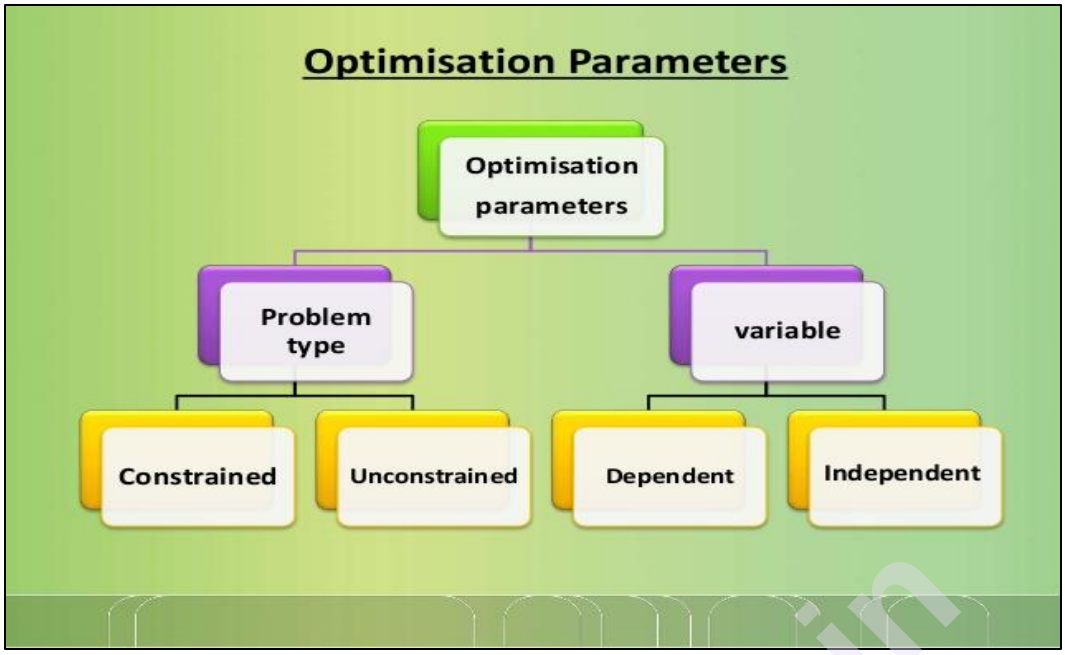
A *nonlinear program* is an optimization problem with an objective function that is an arbitrary nonlinear function of the decision variables, and the constraints can be linear or nonlinear. Nonlinear programs are implemented by the Nonlinear Program class.

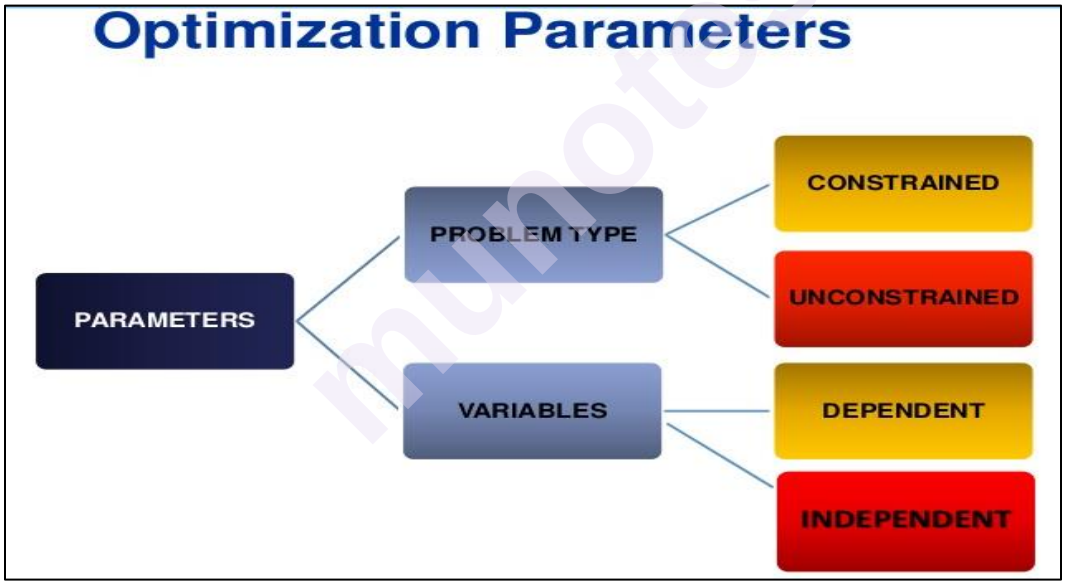
- **<u>Decision Variables</u>**

Finding the optimal values of the decision variables is the goal of solving an optimization model. In the optimization framework, variables are implemented by the Decision Variable class.

Each variable has a name, which may be generated automatically. The Lower Bound and Upper Bound properties specify lower and upper bounds for the values the variable can take. After the solution of the model has been computed, the value property returns the value of the variable in the optimal solution.

munotes.in

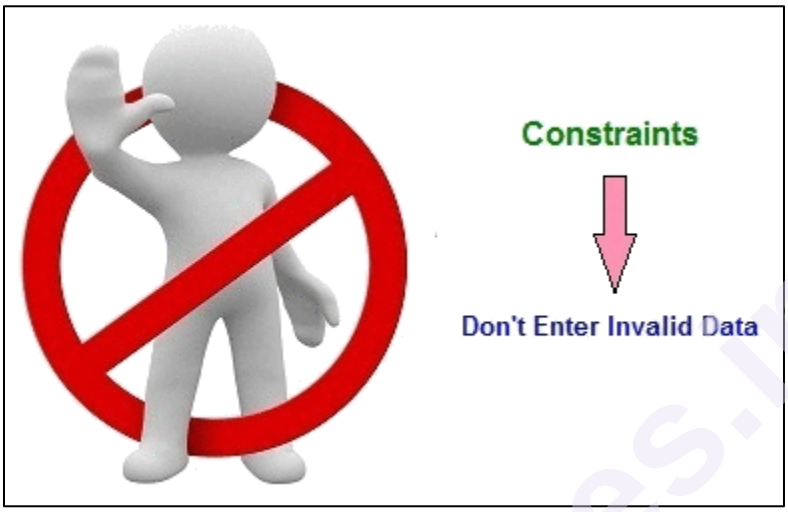




Specific models may have specialized versions of the decision variables. For example, both linear and quadratic programs use variables of type Linear Program Variable, which inherits from Decision Variable and has an extra Cost property that represents the coefficient of the variable in the linear portion of the objective function.

- **<u>Constraints</u>**

Constraints limit the possible values for the decision variables in an optimization model. There are several types of constraints. The classes that implement them all inherit from the constraint class.



Each constraint also has a Name, which may again be generated automatically. The Lower Bound and Upper Bound properties specify lower and upper bounds for the value of the constraint. After the solution of the model has been computed, the Value property returns the value of the constraint in the optimal solution.

munotes.in

Examples of Policy Constraints

- "We will not approve new projects if the projected IRR is less than 20% in three years."
  - Reason given for not pursuing new technology development.
- "If we can't manufacture a component for less than we can buy it, we will outsource that component."
  - International Harvester policy in the 3 years before bankruptcy.
- "We are a metal-stamping company."
  - Reason given for not investing in laser-cutting technology.
- "We can not add components that will increase BOM cost or negatively impact gross margins."
  - Policy that makes decisions blind to impact at the system constraint.

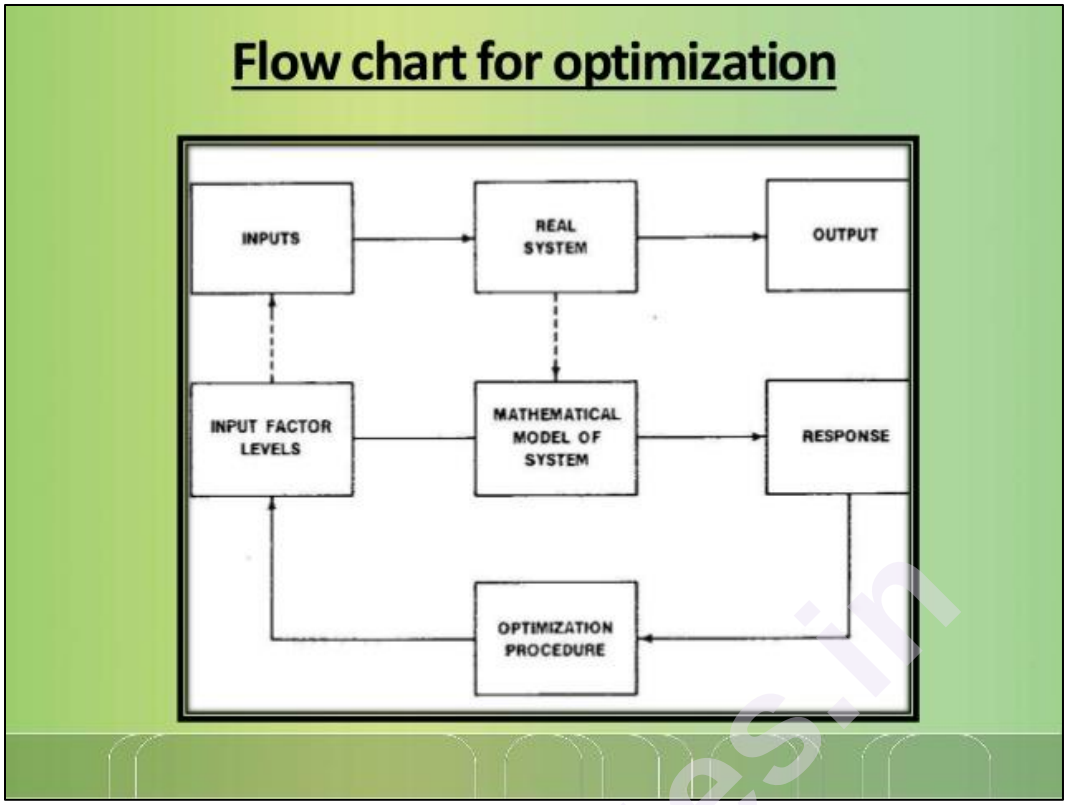
And the most common one of all . . .

- "We strive for efficiency everywhere."

There are two types of constraints: *linear and nonlinear*. Linear constraints express that a linear combination of the decision variables must lie within a certain range. Nonlinear constraints express that the value of some arbitrary function of the decision variables must lie within a certain range.

In mathematics, computer science and operations research, mathematical optimization (alternatively, mathematical programming or simply, optimization) is the selection of a best element (with regard to some criterion) from some set of available alternatives.

In the simplest case, an optimization problem consists of maximizing or minimizing a real function by systematically choosing input values from within an allowed set and computing the value of the function. The generalization of optimization theory and techniques to other formulations comprises a large area of applied mathematics. More generally, optimization includes finding "best available" values of some objective function given a defined domain (or input), including a variety of different types of objective functions and different types of domains.



1A.5. EXERCISE

Question 1. What do you understand by optimization models? Give examples.

Question 2. What are the features of optimization models?

Question 3. Define optimization models. Give examples.

Question 4. Explain the structure and types of optimization models which your own examples.

munotes.in

CHAPTER 1

OPTIMIZATION MODELS

Unit-B

UNIT STRUCTURE

1B.1 Models and Modeling in Operations Research

1B.2 Definitions of Optimization Models

1B.3 Summing Up

1B.4 Exercise

munotes.in

1B.1. MODELS AND MODELING IN OPERATIONS RESEARCH

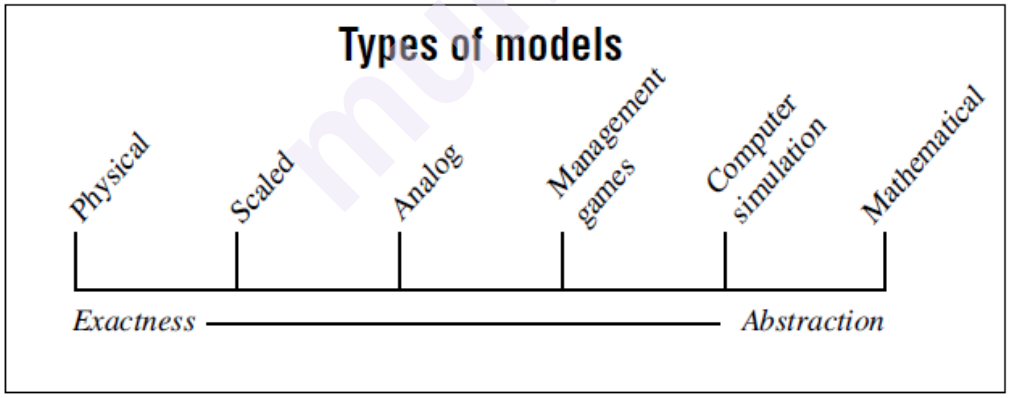
Models

Operations research (OR) is concerned with scientifically deciding how to best design and operate man-machine systems, usually under conditions requiring the allocation of scarce resources.

No matter how OR is defined, the construction and use of models is at its core. Models are representations of real systems. They can be iconic (made to look like the real system), abstract, or somewhere in between.

Iconic models can be full-scale, scaled-down, or scaled-up in size. Sawmill headrig control simulators are full-scale models. A model of the solar system is a scaled-down model, and a teaching model of a wood cell or a water molecule is a scaled-up model. Models can be made of the same material as the system they represent, such as the head rig control simulator, or they can be made of different materials, such as a plastic model of the solar system.

On the other end of the model spectrum are abstract mathematical models. OR professionals often use mathematical models to make simplified representations of complex systems.



Regardless of the type of model used, modeling includes the following steps:

- Defining the problem and gathering data
- Constructing a model of the system
- Deriving a solution

- Testing the model and solution (Is the model valid; that is, does it do what it is designed to do?)
- Implementing the solution.

Modeling in Operations Research (OR)

If the actual system is simple enough to manipulate safely; then OR techniques often are not necessary. For example, managers of the sawmill might wish to extend the hours for loading trucks from 5:00 p.m. until midnight. Since the shipping shed is adjacent to the planer mill, they decide to use the planer mill employees and possibly one or two overtime employees from shipping to extend the loading hours for a trial period of 2 to 3 weeks. At the end of this time, they'll have a good idea of how the extended shipping hours affected the rest of the sawmill operations.

An OR professional also could model this problem, perhaps using a scheduling linear program or a simulation model to examine the effect of extending the shipping hours. However, if the real system can be manipulated without causing too much disruption, it would be preferable to do so rather than to have an OR professional model it. Often, it can take weeks or months just to collect enough data to realistically model a system. In this case, it probably would be less expensive to manipulate the real system.

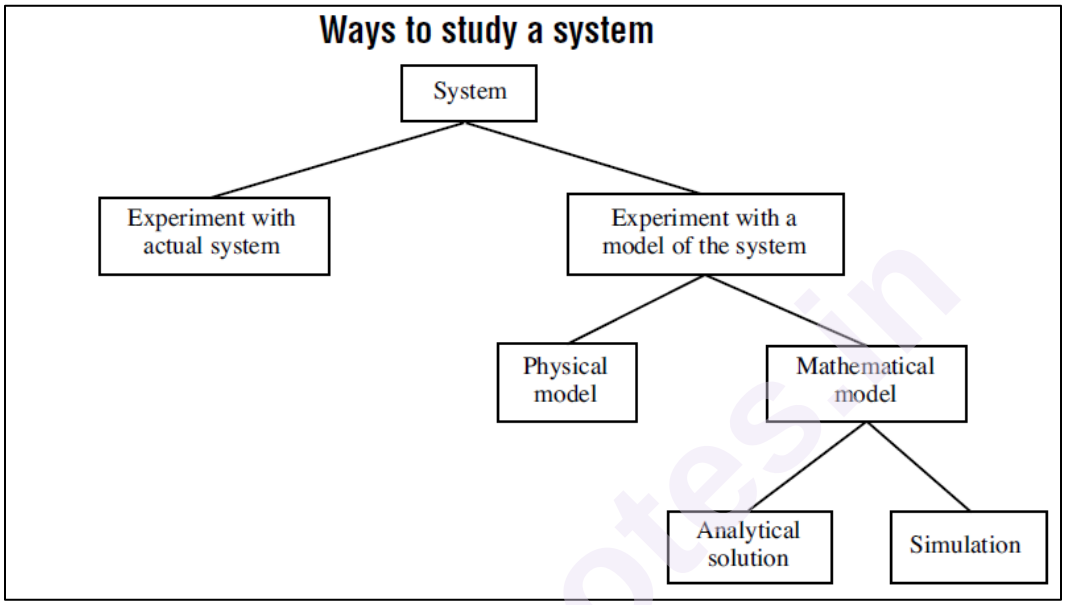
Meteorologists build models to study the weather because they can't experiment with the weather itself. One problem is repeatability. It's impossible to set up repeatable experiments with systems over which the researcher has no control, such as the weather. Meteorologists can observe but not test the system. With a model, however, they can test their theories. They can run the model under different conditions and observe the weather to see whether the model is a good enough (although simplified) representation of the real system. Using the same methods, an experiment using the model is repeatable.

An important, but often overlooked, benefit to modeling is the insight one gets from the process. Often, modelers learn a lot about a system during problem definition and data gathering. An experimenter must organize his or her thoughts to model a system successfully, and this thought

munotes.in

process often reveals previously overlooked insights. Information gained from these new insights can lead to problem resolution even before the model is completed.

An important step in model creation is validation. A model is valid if it does what it was intended to do. The model creator may believe the model is valid, while other users may not. Thus, validation is not as rigid a term as verification or proof and is somewhat subjective.



Operations research experts can study actual systems, but usually use models instead.

What makes a good model?

A simple model is better than a complex one as long as it works as well. A model only needs to perform its intended function to be valid. A model should be easy to understand. It's important to use the most relevant OR tool when constructing a model. A modeler should not try to shape the problem to fit a particular OR method. For example, a linear programming (LP) expert may try to use LP on a problem where there is no optimal solution. Instead, modelers should study the problem and choose the most appropriate OR tool.

For complicated systems, users need to remember that models are only simplified representations. If a user mistakenly considers a complicated model to be correct, he or she may disregard further

study of the real system. Modelers and users of models never should rely only on a model's output and ignore the real system being modeled.

A good model should be easy to modify and update. New information from the real system can be incorporated easily into a well-planned model. A good model usually starts out simple and becomes more complex as the modeler attempts to expand it enough to give meaningful answers.

1B.2. ADVANTAGES AND APPLICATIONS OF OPTIMIZATION MODELS

- **Advantages of Optimization Models**

Most operations research studies involve the construction of a mathematical model. The model is a collection of logical and mathematical relationships that represents aspects of the situation under study. Models describe important relationships between variables; include an objective function with which alternative solutions are evaluated, and constraints that restrict solutions to feasible values.

Although the analyst would hope to study the broad implications of the problem using a systems approach, a model cannot include every aspect of a situation. A model is always an abstraction that is of necessity simpler than the real situation. Elements that are irrelevant or unimportant to the problem are to be ignored; hopefully leaving sufficient detail so that the solution obtained with the model has value with regard to the original problem.

Models must be both tractable, capable of being solved, and valid, representative of the original situation. These dual goals are often contradictory and are not always attainable. It is generally true that the most powerful solution methods can be applied to the simplest, or most abstract, model.

- **Applications of Optimization Models in OR**

Operations Research uses any suitable tools or techniques available. The common frequently used tools/techniques are mathematical procedures, cost analysis, electronic computation. However, operations researchers given special importance to the development and the use of techniques like linear programming, game theory, decision theory, queuing theory, inventory models and simulation. In addition to the above techniques, some other common tools are non-linear

programming, integer programming, dynamic programming, sequencing theory, Markov process, network scheduling (PERT/CPM), symbolic Model, information theory, and value theory. There are many other Operations Research tools/techniques also exists. The brief explanations of some of the above techniques/tools are as follows:

Linear Programming:

This is a constrained optimization technique, which optimize some criterion within some constraints. In Linear programming the objective function (profit, loss or return on investment) and constraints are linear. There are different methods available to solve linear programming.

Game Theory:

This is used for making decisions under conflicting situations where there are one or more players/opponents. In this the motive of the players are dichotomized. The success of one player tends to be at the cost of other players and hence they are in conflict.

Decision Theory:

Decision theory is concerned with making decisions under conditions of complete certainty about the future outcomes and under conditions such that we can make some probability about what will happen in future.

Queuing Theory:

This is used in situations where the queue is formed (for example customers waiting for service, aircrafts waiting for landing, jobs waiting for processing in the computer system, etc). The objective here is minimizing the cost of waiting without increasing the cost of servicing.

Inventory Models:

Inventory model make a decision that minimize total inventory cost. This model successfully reduces the total cost of purchasing, carrying, and out of stock inventory.

munotes.in

Simulation:

Simulation is a procedure that studies a problem by creating a model of the process involved in the problem and then through a series of organized trials and error solutions attempt to determine the best solution. Sometimes this is a difficult/time consuming procedure. Simulation is used when actual experimentation is not feasible or solution of model is not possible.

Non-linear Programming:

This is used when the objective function and the constraints are not linear in nature. Linear relationships may be applied to approximate non-linear constraints but limited to some range, because approximation becomes poorer as the range is extended. Thus, the non-linear programming is used to determine the approximation in which a solution lies and then the solution is obtained using linear methods.

Dynamic Programming:

Dynamic programming is a method of analyzing multistage decision processes. In this each elementary decision depends on those preceding decisions and as well as external factors.

Integer Programming:

If one or more variables of the problem, take integral values only then dynamic programming method is used. For example, number or motor in an organization, number of passenger in an aircraft, number of generators in a power generating plant, etc.

Markov Process:

Markov process permits to predict changes over time information about the behavior of a system is known. This is used in decision making in situations where the various states are defined. The probability from one state to another state is known and depends on the current state and is independent of how we have arrived at that particular state.

munotes.in

Network Scheduling:

This technique is used extensively to plan, schedule, and monitor large projects (for example computer system installation, R & D design, construction, maintenance, etc.). The aim of this technique is minimize trouble spots (such as delays, interruption, production bottlenecks, etc.) by identifying the critical factors. The different activities and their relationships of the entire project are represented diagrammatically with the help of networks and arrows, which is used for identifying critical activities and path.

There are two main types of technique in network scheduling, they are:

- Program Evaluation and Review Technique (PERT) – is used when activities time is not known accurately/ only probabilistic estimate of time is available.
- Critical Path Method (CPM) – is used when activities time is known accurately.

Information Theory:

This analytical process is transferred from the electrical communication field to O.R. field. The objective of this theory is to evaluate the effectiveness of flow of information with a given system. This is used mainly in communication networks but also has indirect influence in simulating the examination of business organizational structure with a view of enhancing flow of information.

Today, almost all fields of business and government utilizing the benefits of Operations Research. There are voluminous of applications of Operations Research. Although it is not feasible to cover all applications of OR in brief, the following are the abbreviated set of typical operations research applications to show how widely these techniques are used today:

Accounting:

- Assigning audit teams effectively
- Credit policy analysis
- Cash flow planning
- Developing standard costs
- Establishing costs for byproducts
- Planning of delinquent account strategy

munotes.in

Construction:

- Project scheduling, monitoring and control
- Determination of proper work force
- Deployment of work force
- Allocation of resources to projects

Facilities Planning:

- Factory location and size decision
- Estimation of number of facilities required
- Hospital planning
- MBA-H2040 Quantitative Techniques for Managers
- International logistic system design
- Transportation loading and unloading
- Warehouse location decision

Finance:

- Building cash management models
- Allocating capital among various alternatives
- Building financial planning models
- Investment analysis
- Portfolio analysis
- Dividend policy making

Manufacturing:

- Inventory control
- Marketing balance projection
- Production scheduling
- Production smoothing

munotes.in

Marketing:

- Advertising budget allocation
- Product introduction timing
- Selection of Product mix
- Deciding most effective packaging alternative

Organizational Behavior / Human Resources:

- Personnel planning
- Recruitment of employees
- Skill balancing
- Training program scheduling
- Designing organizational structure more effectively

Purchasing:

- Optimal buying
- Optimal reordering
- Materials transfer

Research and Development:

- R & D Projects control
- R & D Budget allocation
- Planning of Product introduction

Competitive Strategies

Project Management

munotes.in

1B.3. LET US SUM UP

In this UNIT – B of Chapter 1, you have learnt Models and Modeling in Operations Research, Advantages and Applications of Optimization Models.

Operations Research is relatively a new discipline, which originated in World War II, and became very popular throughout the world. India is one of the few

first countries in the world who started using operations research. Operations Research is used successfully not only in military/army operations but also in business, government and industry. Now a day's operations research is almost used in all the fields.

Proposing a definition to the operations research is a difficult one, because its boundary and content are not fixed. The tools for operations search is provided from the subject's viz. economics, engineering, mathematics, statistics, psychology, etc., which helps to choose possible alternative courses of action. The operations research tool/techniques include linear programming, non-linear programming, dynamic programming, integer programming, Markov process, queuing theory, etc.

Operations Research has a number of applications. Similarly, it has a number of limitations, which is basically related to the time, money, and the problem involves in the model building. Day-by-day operations research gaining acceptance because it improves decision making effectiveness of the managers. Almost all the areas of business use the operations research for decision making.

1B.4. EXERCISE

Question 1. Discuss attributes of a good model and the process of modeling.

Question 2. Explain different modeling techniques of OR.

munotes.in

CHAPTER 2

SEQUENCING MODELS

UNIT STRUCTURE

2.1 Introduction
2.2 Processing n- jobs through Two machine
2.3 Processing n- jobs through Three machines
2.4 Processing n- jobs through m- machines
2.5 Processing Two jobs through m- machines
2.6 Graphical Solution of Two jobs and m- machines Sequencing Problem
2.7 Exercise

munotes.in

2.1. INTRODUCTION

The short-term schedules show an optimal order (sequence) and time in which jobs are processed as well as show timetables for jobs, equipment, people, materials, facilities and all other resources that are needed to support the production plan. The schedules should use resources efficiently to give low costs and high utilisations. Other purpose of scheduling is, minimising customers wait time for a product, meeting promised delivery dates, keeping stock levels low, giving preferred working pattern, minimising waiting time of patients in a hospital for different types of tests and so on.

The general scheduling or sequencing problem may be described as: Let there be n jobs to be performed one at a time on each of m machines. The sequence (order) of the machines in which each job should be performed is given. The actual or expected time required by the jobs on each of the machines is also given. The general sequencing problem, therefore, is to find the sequence out of $(n!)^m$ possible sequences which minimise the total elapsed time between the start of the job in the first machine and the completion of the last job on the last machine.

In particular, if n = 3 and m= 3, then total number of possible sequences will be $(3!)^3 = 216$. Theoretically, it may be possible to find optimum sequence but it will require a big computational time. Thus, one should adopt sequencing technique.

To find optimum sequence we first calculate the total elapsed time for each of the possible sequences. As stated earlier, even if values of m and n are very small, it is difficult to get the desired sequence with total minimum elapsed time. However, due to certain rules designed by Johnson, the task of determining an optimum sequence has become quite easy.

NOTATIONS, TERMINOLOGY AND ASSUMPTIONS

Notations

t_{ij} = Processing time (time required) for job i on machine j.

T = Total elapsed time for processing all the jobs. This includes idle time, if any.

I_{ij} = Idle time on machine j from the end of job (j - 1) to the start of job i.

Terminology

- **Number of Machines:** The number of machines refer to the number of service facilities through which a job must pass before it is assumed to be completed.
- **Processing Time:** It is the time required by a job on each machine.
- **Processing Order:** It refers to the order (sequence) in which machines are required for completing the job.
- **Idle Time on a Machine:** It is the time for which a machine does not have a job to process, i.e. idle time from the end of job (i-1) to the start of job i.
- **Total Elapsed Time:** It is the time interval between starting the first job and completing the last job including the idle time (if any) in a particular order by the given set of machines.
- **No Passing Rule:** It refers to the rule of maintaining the order in which jobs are to be processed on given machines. For example, if n jobs are to be processed on two

machines M_1 and M_2 in the order M_1 M_2, then each job should go to machine M_1 first and then to M_2.

Assumptions

1. The processing time on different machines are exactly known and are independent of the order of the jobs in which they are to be processed.
2. The time taken by the job in moving from one machine to another is negligible.
3. Once a job has begun on a machine, it must be completed before another job can begin on the same machine.
4. All jobs are known and are ready for processing before the period under consideration begins.
5. Only one job can be processed on a given machine at a time.
6. Machines to be used are of different types.
7. The order of completion of jobs are independent of the sequence of jobs.

2.2. PROCESSING n JOBS THROUGH TWO MACHINES

Let there be n jobs, each of which is to be processed through two machines, M_1 and M_2 in the order M_1 M_2, i.e. each job has to be passed through the same sequence of operations. In other words, a job is assigned on machine M_1 first and after it has been completely processed on machine M_1, it is assigned to machine M_2. If the machine M_2 is not free at the moment for processing the same job, then the job has to wait in waiting line for its turn on machine M_2, i.e. passing is not allowed.

Since passing is not allowed, therefore, machine M_1 will remain busy in processing all the n jobs one-by-one which machine M_2 may remain idle time of the second machine. This can be achieved only by determining sequence of n jobs which are to be processed on two machines M_1 and M_2. The procedure suggested by Johnson for determining the optimal sequence can be summarised as follows:

The Algorithm

Step 1 List the jobs along with their processing times on each machine in a table as shown below:

Processing Time on Machine	Job Number			
	1	2	3	n
M_1	t_{11}	t_{12}	t_{13}	t_{1n}
M_2	t_{21}	t_{22}	t_{23}	t_{2n}

Step 2 Examine the columns for processing times on machines M_1 and M_2, and find the smallest processing time in each column, i.e. find out, min. (t_{1j}, t_{2j}) for all j.

Step 3 (a) If the smallest processing time is on machine M_1, then schedule the job as early as possible without moving jobs already schedules, i.e. place the job in the first available position in the sequence. If the processing time is on machine M_2, then schedule the job as late as possible without moving any jobs already scheduled, i.e. place the job in the last available position in the sequence.

(b) If there is a tie in selecting the minimum of all the processing times, then there may be three situations:

a. Minimum among all processing times is same for the machine i.e. min (t_{1j}, t_{2j}) = t_{2k} = t_{2r}, then process the kth job first and the rth job last.

b. If the tie for the minimum occurs among processing times t_{1j} on machine M_1 only, then select the job corresponding to the smallest job subscript first.

c. If the tie for the minimum occurs among processing times t_{2j} on machine M_2, then select the job corresponding to the largest job corresponding to the largest job subscript last.

Step 4 Remove the assigned jobs from the table. If the table is empty, stop and go to Step 5. Otherwise, got to Step 2.

Step 5 Calculate idle time for machines M_1 and M_2:

a. Idle time for machine M_1 = (Total Elapsed Time) - (Time when the last job in a sequence finishes on machine M_1)

b. Idle time for machine M_2 = Time at which the first job in a sequence finishes on machine M_1 + $\sum_{j=2}^{n}$ {(Time when the jth job in a sequence starts on machine M_2) - (Time when the (j - 1)th job in a sequence finishes on machine M_2)}.

Step 6 The total elapsed time to process all jobs through two machines is given by

Total Elapsed time = Time when the nth job in a sequence finishes on Machine M_2.

= $\sum_{j=1}^{n} M2j + \sum_{j=1}^{n} I2j$

where M_{2j} = Time required for processing j^{th} job on machine M_2.

I_{2j} = Time for which machine M_2 remains idle after processing (j - 1)th job and before starting work in j^{th} job.

EXAMPLE 1

Find the sequence that minimises the total elapsed time required to complete the following tasks on two machines:

Task	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
Machine I	**2**	**5**	**4**	**9**	**6**	**8**	**7**	**5**	**4**
Machine II	**6**	**8**	**7**	**4**	**3**	**9**	**3**	**8**	**11**

Solution:

The smallest processing time between the two machines is 2 which corresponds to task A on Machine I. Thus, task A is scheduled as early as possible to give the sequence as shown below:

A								

After the task A has been set for processing first, we are left with 8 tasks and their processing times as given below:

Task	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
Machine I	**5**	**4**	**9**	**6**	**8**	**7**	**5**	**4**
Machine II	**8**	**7**	**4**	**3**	**9**	**3**	**8**	**11**

The minimum processing time in this reduced problem is 3 which corresponds to task E and G both on machine II. Since the corresponding processing time of task E on machine I is less

than the corresponding processing time of task G on machine I, therefore, task E will be scheduled in the last and task G shall be scheduled before it. Tasks E and G will not be considered further. Thus, current partial sequence of scheduling tasks becomes:

A							G	E

A set of processing times now gets reduced to:

Task	**B**	**C**	**D**	**F**	**H**	**I**
Machine I	**5**	**4**	**9**	**8**	**5**	**4**
Machine II	**8**	**7**	**4**	**9**	**8**	**11**

The smallest processing time in this reduced problem is 4, which corresponds to task C and I on machine I and to task D on machine II. Thus task C will be placed in the second sequence cell and task I in the third sequence cell and task D in the sequence cell before task G. The entries of the partial sequence are now:

A	C	I				D	G	E

The set of processing time now gets reduced as follows:

Task	**B**	**F**	**H**
Machine I	**5**	**8**	**5**
Machine II	**8**	**9**	**8**

the smallest processing time in this reduced problem is 5, which corresponds to tasks B and H both on machine I. Since the corresponding processing times of B and h on machine II is same, therefore, either of these two tasks can be placed in fourth and fifth sequence cells. Thus, it indicates an alternative optimal sequence. the optimal sequences are, therefore, given below:

A	C	I	B	H	F	D	G	E

A	C	I	H	B	F	D	G	E

The minimum elapsed time for machines I and II is calculated as shown in Table 1.

Task Sequence	Machine I		Machine II	
	Time In	Time Out	Time In	Time Out
A	0	2	2	8
C	2	6	8	15
I	6	10	15	26
B	10	15	26	34
H	15	20	34	42
F	20	28	42	51
D	28	37	51	55
G	37	44	55	58
E	44	50	58	61

In table 1, the minimum elapsed time, i.e. time from start of task A to completion of last task E is 61 hours. During this time the machine I remains idle for 61 - 50 = 11 hours. The idle time for machine II is equal to the time at which the first task A in the sequence finishes on machine I plus the last task E in the sequence starts on machine II minus the last but one task G finishes on machine II, i.e. 2 + 58 - 58 = 2 hours.

munotes.in

2.3. PROCESSING n JOBS THROUGH THREE MACHINES

Johnson provides an extension of his procedure to the case in which there are three instead of two machines. Each job is to be processed through three machines M_1, M_2 and M_3. The list of jobs with their processing times is given below. An optimal solution to this problem can be obtained if either or both of the following conditions hold good.

Processing time on Machine	Job Number			
	1	2	3	4
M_1	t_{11}	t_{12}	t_{13}	t_{1n}
M_2	t_{21}	t_{22}	t_{23}	t_{2n}
M_3	t_{31}	t_{32}	t_{33}	t_{3n}

1. The minimum processing time on machine M_1 is at least as great as the maximum processing time on machine M_2, that is,

$$\min t_{1j} \geq \max t_{1j}, \qquad j = 1, 2, 3, ...n$$

2. The minimum processing time on machine M_3 is at least as great as the maximum processing time on machine M_2, that is

$$\min t_{3j} \geq \max t_{2j}, \qquad j = 1, 2, 3, ...n$$

If either or both the above conditions hold good, then the steps of the algorithm can be summarised in the following steps:

THE ALGORITHM

Step 1: Examine processing times of given jobs on all three machines and if either one or both the above conditions hold, then go to step 2, otherwise the algorithm fails.

Step 2: Introduce two fictitious machines, say G and H with corresponding processing times given by

$$\text{i.} \quad t_{Gj} = t_{1j} + t_{2j}, \; j = 1, 2, 3, ...n.$$

that is, processing time on machine G is the sum of the processing times on machines M_1 and M_2, and

$$\text{ii.} \quad t_{Hj} = t_{2j} + t_{3j}, \; j = 1, 2, 3, ...n.$$

that is, processing time on machine H is the sum of the processing times on machines M_2 and M_3.

Step 3: Determine the optimal sequence of jobs for this n-job, two machine equivalent sequencing problem with the prescribed ordering GH in the same way as discussed earlier.

EXAMPLE

Find the sequence that minimises the total time required in performing the following jobs on three machines in the order ABC. Processing times (in hours) are given in the following table:

Job	1	2	3	4	5
Machine A	8	10	6	7	11
Machine B	5	6	2	3	4

munotes.in

Machine C	4	9	8	6	5

Solution: Here, min (t_{Aj}) = 6; Min (t_{Cj}) = 4; max (t_{Bj}) = 6. Since min (t_{Aj}) ≥ (t_{Bj}) for all j is satisfied, the given problem can be converted into a problem of 5 jobs and two machines. The processing time on two fictitious machines G and H can be determined by the following relationships:

$t_{Gj} = t_{Aj} + t_{Bj}$, j = 1, 2, 3, ...n.

and $t_{Hj} = t_{Bj} + t_{Cj}$, j = 1, 2, 3, ...n.

The processing times for the new problem are given below:

Job	1	2	3	4	5
Machine G	8 + 5 = 13	10 + 6 = 16	6 + 2 = 8	7 + 3 = 10	11 + 4 = 15
Machine H	5 + 4 = 9	6 + 9 = 15	2 + 8 = 10	3 + 6 = 9	4 + 5 = 9

When the algorithm described for n jobs on two machines is applied to this problem, the optimal sequence so obtained is given by

3	2	5	1	4

The total minimum elapsed time is given in Table 1:

Job Sequence	Machine A		Machine B		Machine C	
	Time In	Time Out	Time In	Time Out	Time In	Time Out
3	0	6	6	8	8	16
2	6	16	16	22	22	31
5	16	27	27	31	31	36
1	27	35	35	40	40	44
4	35	42	42	45	45	51

Table 1 indicates that the minimum total elapsed time is 51 hours. The idle time for machines A, B and C is 9 (=51 - 42) hours, 6 (= 51 - 45) hours and 9 (= 8 - 0) + (45 - 44) hours, respectively.

2.4. PROCESSING n JOBS THROUGH m MACHINES

Let there be n jobs, each of which is to be processed through m machines, say M_1, M_2, ...M_m in the order M_1, M_2, ...M_m. The optimal solution to this problem can be obtained if either or both of the following conditions hold good.

(a) Min $\{t_{1j}\}$ ≥ Max $\{ t_{1j}\}$; j = 2, 3, ... m - 1

and or (b) Min $\{t_{mj}\}$ ≥ Max $\{ t_{ij}\}$; j = 2, 3, ... m - 1

that is, the minimum processing time on machines M_1 and M_m is as great as the maximum processing time on any of the remaining (m - 1) machines.

If either or both these conditions hold good, then the steps of the algorithm can be summarised in the following steps:

Step 1: Find, Min $\{t_{1j}\}$, Min $\{t_{mj}\}$ and max $\{t_{ij}\}$and verify above conditions. If either or both the conditions mentioned above hold, then go to step 2. Otherwise the algorithm fails.

Step 2: Convert m-machine problem into 2-machine problem by introducing two fictitious machines, say

(i) $t_{Gj} = t_{1j} + t_{2j} + t_{3j} + ... + t_{m-1j}$ j = 1, 2, 3, ...n.

i.e. processing time of n-jobs on machine G is the sum of the processing times on Machines M_1, M_2 .....M_{m-1j}

(ii) $t_{Hj} = t_{2j} + t_{3j} + t_{4j} + ... + t_{mj}$ j = 1, 2, 3, ...n.

i.e. processing time of n-jobs on machine H is the sum of the processing times on Machines M_1, M_2 .....$M_{mj.}$

Step 3: The new processing times so obtained can now be used for solving n-job, two machines equivalent sequencing problem with the prescribed ordering HG in the same way as $t_{2j} + t_{3j} + ... + t_{m-1j} = k$ (constant)

for all j = 1, 2, 3, ... m - 1, then the optimal sequence can be obtained for n-jobs and two machines M_1 and M_m in the order M_1 M_m as usual.

2. If $t_{1j} = t_{mj}$ and $t_{Gj} = t_{Hj,}$ for all j = 1, 2, 3, ...n, then total number of optimal sequences will be n and total minimum elapsed time in these cases would also be the same.

3. The method described above for solving n-jobs and m-machines sequencing problem is not a general method. It is applicable only to certain problems where the minimum cost (or time) of processing the jobs through first and/or last machine is more than or equal to the cost (or time) of processing the jobs through remaining machines.

EXAMPLES

1. Find an optimal sequence for the following sequencing problems of four jobs and five machines when passing is not allowed of which processing time (in hours) is given below:

Job	Machines				
	M_1	M_2	M_3	M_4	M_5
A	7	5	2	3	9
B	6	6	4	5	10
C	5	4	5	6	8
D	8	3	3	2	6

Also find the total elapsed time.

Solution: Here,

Min $(t_{M1,\,j}) = 5 = t_{M1,\,C}$

Min $(t_{M5,\,j}) = 6 = t_{M5,\,D}$

and Max $\{t_{M2,\,j}, t_{M3,\,j}, t_{M4,\,j}\} = \{6, 5, 6\}$ respectively.

Since the condition of Min $(t_{M5,\,j}) \geq$ Max $\{ t_{M2,\,j}, t_{M3,\,j}, t_{M4,\,j}\}$ is satisfied, therefore the given problem can be converted into a four jobs and two machines problem as G and H. The processing times of four jobs denoted by t_{Gj} and t_{Hj} on G and H, respectively are as follows:

Job	A	B	C	D
Machine G	17	21	20	16
Machine H	19	25	23	14

where $t_{Gj} = \sum_{i=1}^{m-1} tij$ and $t_{Hj} = \sum_{i=2}^{m} tij$.

Now using the optimal sequence algorithm, the following optimal sequence can be obtained.

A	C	B	D

The total elapsed time corresponding to the optimal sequence can be calculated as shown in Table 1, using the individual processing times given in the original problem.

Table 1 shows that the minimum total elapsed time is 51 hours. The idle time for machines M_1, M_2, M_3, M_4 and M_5 is 25, 33, 37 and 18 hrs respectively.

Table 1 Minimum Elapsed Time

Job Sequence	Machine				
	M_1	M_2	M_3	M_4	M_5
A	0 - 7	7 - 12	12 - 14	14 - 17	16 - 26
B	7 - 12	12 - 18	16- 21	21 - 27	27 - 35
C	12 - 18	18 - 24	24 - 28	28 - 33	35 - 45
D	18 - 26	26 - 29	29 - 32	33 - 35	45 - 51

2. Solve the following sequencing problem giving an optimal solution when passing is not allowed.

Machine	Job				
	A	B	C	D	E
M_1	11	13	9	16	17
M_2	4	3	5	2	6
M_3	6	7	5	8	4
M_4	15	8	13	9	11

Solution: From the data of the problem it is observed that

Min $(t_{M1, j}) = 9 = t_{M1, C}$

Min $(t_{M4, j}) = 8 = t_{M4, B}$

and Max $\{t_{M2, j}\} = 6 = , t_{M2, E;}$ Max $\{ t_{M3, j}\} = 8 = , t_{M2, D}$.

Since both the conditions

Min $(t_{M1, j}) \geq$ Max $\{ t_{M2, j;} t_{M3, j}\}$; j = 1, 2, ..., 5

are satisfied, therefore given problem can be converted into a 5-jobs and 2-machine problem as G and H.

Further, it may be noted that, $t_{M2, j} + t_{M3, j} = 10$ (a fixed constant) for all j (j = 1, 2, ..., 5).

Thus the given problem is reduced to a problem of solving 5-jobs through 2-machines M_1 and M_4 in the order M_1 M_4. This means machines M_2 and M_4 will have no effect on the optimality of the sequences.

The processing times of 5 jobs on machine M_1 and M_4 is as follows:

Job	A	B	C	D	E
Machine M_1	11	13	9	16	17
Machine M_4	15	8	13	9	11

Now using the algorithm described earlier, the optimal sequence so obtained as follows:

C	A	E	D	B

The total elapsed time corresponding to the optimal sequence is 83 hours as shown in table 1, using the individual processing times given in the original problem:

Table 1 Minimum Total Elapsed Time

Job Sequence	Machine				
	M_1	M_2	M_3	M_4	
C	0 - 9	9 - 14	14 - 19	19 - 32	
A	9 - 20	20 - 24	24 - 30	32 - 45	
E	29 - 36	36 - 42	42 - 46	46 - 57	
D	36 - 52	52 - 54	54 - 62	62 - 71	
B	52 - 65	65 - 68	68 - 75	75 - 83	

munotes.in

2.5. PROCESSING TWO JOBS THROUGH m MACHINES

Let there be two jobs A and B each of which is to processed on m machines say M_1, M_2, ..., M_m in two different orders. The technological ordering of each of the two jobs through m machines is known in advance. Such ordering may not be same for both the jobs. The exact or expected processing times on the given machines are known. Each machine can perform only one job at a time. The objective is to determine an optimal sequence of processing the jobs so as to minimise total elapsed time.

The optimal sequence in this case can be obtained by using graph. The procedure can be illustrated by taking examples.

Example 1: Use the graphical method to minimise the time needed to process the following jobs on the machines shown, i.e. each machine finds the job which should be done first. Also calculate the total elapsed time to complete both jobs.

Machine

Job 1 {Sequence:	A	B	C	D	E
Time (hrs)	3	4	2	6	2

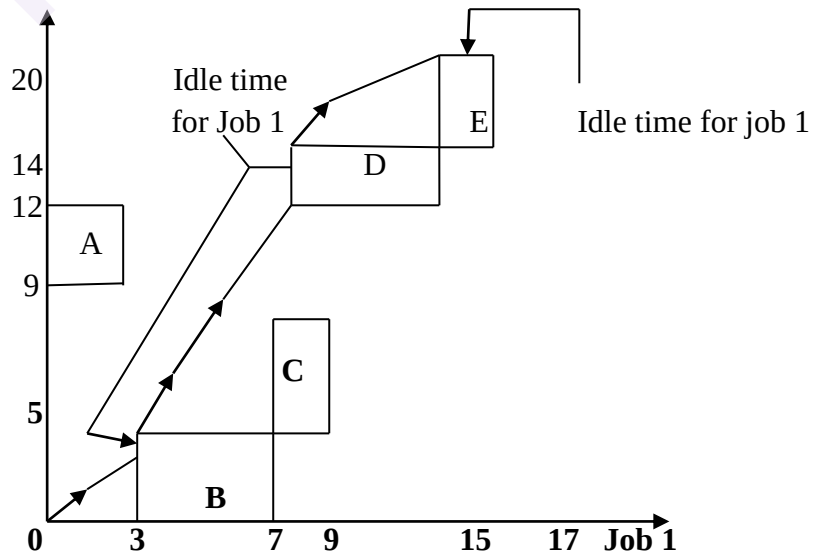
Machine

Job 2 {Sequence:	A	B	C	D	E
Time (hrs)	5	4	3	2	6

Solution

The solution procedure for solving the above problem can be summarised in the following steps:

1. Draw the set of axes at right angle to each other where x-axis represents the processing time of job 1 on different machines while job 2 remains idle and y-axis represents processing time of job 2 while job 2 remain idle.
2. Mark the processing times for jobs 1 and 2 on x-axis and y-axis, respectively according to the given order of machines as shown in the figure:



2.6. GRAPHICAL SOLUTION OF TWO JOBS AND m- MACHINES SEQUENCING PROBLEM

For example, machine A takes 3 hours for job 1 and 3 hours for job 2. Construct the rectangle for machine A as shown in above Figure. Similarly, construct other rectangles for machines B, C, D and E.

3. Construct various blocks starting from the origin by pairing the same machine until a point marked 'finished' is obtained.

4. Draw a line starting from origin to the point marked 'finish' by moving horizontally, vertically and diagonally along a line which makes an angle of 45° with the horizontal axis. Moving horizontally along this line indicates that first job is under process while second job is idle. Similarly, moving vertically along this line indicates that the second job is under process while first job is idle. The diagonal movement along this line shows that both the jobs are under process simultaneously.

Since simultaneous processing of both jobs on a machine is not possible, therefore, diagonal movement is not allowed. In other words, diagonal movement through rectangle areas is not allowed.

5. An optimal path is one that minimises the idle time for both the jobs. Thus, we must choose the path on which diagonal movement is maximum as shown in the above figure.

6. Total elapsed time is obtained by adding the idle time for either job to the processing time for that job. In this example, the idle time for the chosen path is found to be 5 hrs and 2 hrs for jobs 1 and 2, respectively. The total elapsed time is calculated as follows:

Elapsed time, Job 1 = Processing time of Job 1 + Idle time for Job 1

= 17 + (2 + 3) = 22 hrs.

Elapsed time, Job 2 = Processing time of Job 2 + Idle time for Job 2

= 22 + (17-15) = 24 hours.

munotes.in

2.7. EXERCISE

1. Explain the four elements that characterise a sequencing problem.

2. Explain the principal assumptions made while dealing with sequencing problems.

3. Give Johnson's procedure for determining an optimal sequence for processing n items on two machines. Give justification of the rule used in the procedure.

4. What is no passing rule in a sequencing algorithm? Explain the principal assumptions made while dealing with sequencing problems.

5. Give three different examples of sequencing problems from your daily life.

6. We have five jobs, each of which must be processed on the two machines A and B in the order AB. Processing times in hours are given in the table below:

Job	1	2	3	4	5
Machine A	5	1	9	3	10
Machine B	2	6	7	8	4

7. We have 5 jobs each of which must go through the machines A, B and C in the order ABC. Processing times (in hours) is as follows:

Job	1	2	3	4	5
Machine A	5	7	6	9	5
Machine B	2	1	4	5	3
Machine C	3	7	5	6	7

8. What do you understand by the following terms in the context of sequence of jobs:
1. Job arrival pattern 2. Number of machines 3. The flow pattern in the shop
4. the criteria of evaluating the performance of a schedule.

9. Find an optimal sequence for the following sequencing problem of four jobs and five machines (when passing is not allowed) of which processing time (in hrs) is as follows:

Job	1	2	3	4
Machine M_1	6	5	4	7
Machine M_2	4	5	3	2
Machine M_3	1	3	4	2
Machine M_4	2	4	5	1
Machine M_5	8	9	7	5

Also find the total elapsed time.

10. Two jobs are to be processed on four machines A, B, C and D. The technological order for these jobs on machines is as follows:

Job 1: A B C D
Job 2: D B A C

Processing times are given in the following table:

	Machines			
	A	B	C	D
Job 1	4	6	7	3
Job 2	4	7	5	8

Find the optimal sequence of jobs on each of the machines.

munotes.in

CHAPTER 3

REPLACEMENT MODELS

UNIT STRUCTURE

3.1 Introduction
3.2 Types of Failure
3.3 Replacement of items whose efficiency deteriorates with time
3.4 Replacement of items that fail completely
3.5 Exercise

munotes.in

3.1. INTRODUCTION

The problem of replacement is felt when the job performing units such as men, machines, equipments, parts etc. become less effective or useless due to either sudden or gradual deterioration in their efficiency, failure or breakdown. By replacing them with new ones at frequent intervals, maintenance and other overhead costs can be reduced. However, such replacements would increase the need of capital cost for new ones.

For example,

1. A vehicle tends to wear out with time due to constant use. More money needs to be spent on it on account of increased repair and operating cost. A stage comes when it becomes uneconomical to maintain the vehicle and it is better to replace it with a new one. Here the replacement decision may be taken to balance the increasing maintenance cost with the decreasing money value of the vehicle as time passes.
2. In case of highway tube lights where time of failure is not predictable for individual tubes, replacement is done only after individual failure. However, it may be economical to replace such tubes on a schedule basis before their failure. Here the replacement decision may be taken to balance between the wasted life of a tube before failure and cost incurred when a tube fails completely during service.

Thus, the basic problem in such situations is to formulate a replacement policy to determine an age (or period) at which the replacement of the given machinery/equipment is most economical keeping in view all possible alternatives.

In this chapter, we shall discuss the replacement policies in the context of the following three types of replacement situations:

1. Items such as machines, vehicles, tyres etc. whose efficiency deteriorates with age due to constant use and which need increased operating and maintenance costs. In such cases deterioration level is predictable and is represented by a. Increased maintenance/operational cost, b. its waste or scrap value and damage to item and safety risk.
2. Items such as light bulbs and tubes, electric motors, radio, television parts etc. which do not give any indication of deteriorations with time but fail all of a sudden and become completely useless. Such cases require an anticipation of failures to specify the probability of failure for any future time period. With this probability distribution and the cost information, it is desired to formulate optimal, replacement policy to balance the wasted life of an item replaced before failure against the costs incurred when the item fail in service.
3. The existing working staff in an organisation gradually reduces due to retirement, death, retrenchment and other reasons.

3.2. TYPES OF FAILURE

The term 'failure' here will be discussed in the context of replacement decisions. There are two types of failures: 1. Gradual failure and 2. Sudden failure.

Gradual Failure

Gradual Failure is progressive in nature. That is, as the life of an item increases its operational efficiency also deteriorates resulting in

- increased running (maintenance and operating) costs.
- decrease in its productivity.
- decrease in the resale or salvage value.

Mechanical items like pistons, rings, bearings etc., and automobile types fall under this category.

Sudden Failure

This type of failure occurs in items after some period of giving desired service rather than deterioration while in service. The period of desired service is not constant but follows some frequency distribution which may be progressive, retrogressive or random in nature.

1. Progressive Failure: If the probability of failure of an item increases with the increase in its life, then such failure is called progressive failure. For example, light bulbs and tubes fail progressively.

2. Retrogressive Failure: If the probability of failure in the beginning of the life of an item is more but as time passes the chances of its failure becomes less, then such failure is said to be retrogressive.

3. Random Failure: In this type of failure, the constant probability of failure is associated with items that fail from random cases such as physical shocks, not related to age. For example, vacuum tubes in air-born equipment have been found to fail at a rate independent of the age of the tube.

3.3. REPLACEMENT OF ITEMS WHOSE EFFICIENCY DETERIORATES WITH TIME

When operational efficiency of an item deteriorates with time (gradual failure), it is economical to replace the same with a new one. For example, the maintenance cost of a machine increases with time and a stage is reached when it may not be economical to allow machine to continue in the system. Besides, there could be a number of alternative choices and one may like to compare available alternatives on the basis of the running costs (average maintenance and operating costs) involved. In this section, we shall discuss various techniques for making such comparisons under different conditions. While making such comparisons it is assumed that suitable expressions for running costs are available.

Model 1 Replacement Policy for items Whose Running Cost Increases with Time and Value of Money Remains Constant During a Period

Theorem 1

The cost of maintenance of a machine is given as a function increasing with time and its scrap value is constant.

a. If time is measured continuously, then the average annual cost will be minimised by replacing the machine when the average cost to date becomes equal to the current maintenance cost.

b. If time is measured in discrete units, then the average annual cost will be minimised by replacing the machine when the next period's maintenance cost becomes greater than the current average cost.

PROOF: The aim here is to determine the optimal replacement age of an equipment whose running cost increases with time and the value of money remains constant (i.e. value is not counted) during that period.

Let C = capital or purchase cost of new equipment

S = Scrap (or salvage) value of the equipment at the end of t years

R(t) = running cost of equipment for the year t

n = replacement age of the equipment.

i. When time 't' is a continuous variable: If the equipment is used for t years, then the total cost incurred over this period is given by

TC = Capital (or purchase) cost - Scrap value at the end of t years + Running cost for t years.

$= C - S + \int_0^n R(t)dt$

Therefore, the average cost per unit time incurred over the period of n years is:

$ATC_n = \frac{1}{n}\left\{C - S + \int_0^n R(t)dt\right\}$ ---------- (1)

To obtain optimal value n for which ATC_n is minimum, differentiate ATC_n with respect to n, and set the first derivative equal to zero. That is, for minimum of ATC_n.

$\frac{d}{dn}\{ATC_n\} = -\frac{1}{n2}\{C - S\} + \frac{R(n)}{n} - \frac{1}{n2}\int_0^n R(t)dt = 0$

or $R(n) = \frac{1}{n}\left\{C - S + \int_0^n R(t)dt\right\}$, $n \neq 0$ ---------- (2)

$R(n) = ATC_n$

Hence, the following replacement policy can be derived with help of Eq. 2.

Policy

Replace the equipment when the average annual cost for n years becomes equal to the current /annual running cost. That is,

$$R(n) = \frac{1}{n}\left\{C - S + \int_0^n R(t)dt\right\}$$

b. When time 't' is a discrete variable: The average cost incurred over the period n is given by

$ATC_n = \frac{1}{n}\left\{C - S + \sum_{t=0}^{n} R(t)\right\}$ ----------(3)

If C - S and $\sum_{t=0}^{n} R(t)$ are assumed to be monotonically decreasing and increasing respectively, then there will exist a value of n for which ATC_n is minimum. Thus we shall have inequalities

$ATC_{n-1} > ATC_n < ATC_{n+1}$

or $ATC_{n-1} - ATC_n > 0$

and $ATC_{n+1} - ATC_n > 0$

Eq. (3) for period n + 1, we get

$ATC_n = \frac{1}{n+1}\left\{C - S + \sum_{t=1}^{n=0} R(t)\right\} = \frac{1}{n+1}\left\{C - S + \sum_{t=1}^{n=0} R(t) + R(n+1)\right\}$

$= \frac{n}{n+1} \frac{C - S + \sum_{t=0}^{n} R(t)}{n} + \frac{R(n+1)}{n+1} = \frac{n}{n+1} . ATC_n + \frac{R(n+1)}{n+1}$

Therefore,

$ATC_{n+1} - ATC_n = \frac{n}{n+1} ATC_n + \frac{A(n+1)}{n+1} - ATC_n = \frac{R(n+1)}{n+1} + ATC_n [\frac{n}{n+1} - 1] = \frac{R(n+1)}{n+1} - \frac{ATCn}{n+1}$

Since $ATC_{n+1} - ATC_n > 0$, we get

$\frac{R(n+1)}{n+1} - \frac{ATCn}{n+1} > 0$

$R(n + 1) - ATC_n > 0$ or $R(n + 1) > ATC_n$ ---------- (4)

Similarly,

$ATC_{n-1} - ATC_n > 0$, implies that $R(n + 1) < ATC_{n-1}$. This provides the following replacement policy.

Policy 1: If the next year, running cost R(n + 1) is more than average cost of nth year, ATC_n, then it is economical to replace at the end of n years. That is

$R(n + 1) > \left\{\frac{1}{n} C - S + \sum_{t=0}^{n} R(t)\right\}$

Policy 2: If the present year's running cost is less than the previous year's average cost, ATC_{n-1}, then do not replace. That is

$R(n) < \frac{1}{n-1}\left\{C - S + \sum_{t=0}^{n-1} R(t)\right\}$

EXAMPLE

A firm is considering replacement of a machine, whose cost price is Rs. 12, 200 and scrap value is Rs. 200. The running (maintenance and operating) costs are found from experience to be as follows:

Year	1	2	3	4	5	6	7	8
Running cost (Rs.)	200	500	800	1200	1800	2500	3200	4000

When should the machine be replaced?

Solution: We are given the running cost, R(n), the scrape value S - Rs. 200 and the cost of the machine, C = Rs. 12, 200. In order to determine the optimal time n when the machine should be replaced, first we calculate average cost per year during the life of the machine as shown in table below:

TABLE

Year of Service n	Running Cost (Rs.) R(n)	Cumulative Running Cost (Rs.) Σ R(n)	Depreciation Cost (Rs.) C - S	Total Cost (Rs.) TC	Average Cost (Rs.) ATC_n
1	2	3	4	5 = 3 + 4	6 = 5/1
1	200	200	12000	12, 200	12, 200
2	500	700	12000	12, 700	6350
3	800	1500	12000	13, 500	4500
4	1200	2700	12000	14, 700	3675
5	1800	4500	12000	16, 500	3300

6	2500	7000	12000	19, 000	3167
7	3200	10200	12000	22, 200	3171
8	4000	14200	12000	26, 200	3275

From the table, it may be noted that the average cost per year, ATC_n is minimum in the sixth year (Rs. 3, 167). Also the average cost in the seventh year (Rs. 3171) is more than the cost in the sixth year. Hence, the machine should be replaced after every six years.

(2) The data collected in running a machine, the cost of which is Rs. 60, 000 are given below:

Year	1	2	3	4	5
Resale Value (Rs.)	42, 000	30, 000	20, 400	14, 400	9, 650
Cost of Spares (Rs.)	4, 000	4, 270	4, 880	5, 700	6, 800
Cost of Labour (Rs.)	14, 000	16, 000	18, 000	21, 000	25, 000

Determine the optimum period for replacement of the machine.

Solution: The costs of spares and labour together determine running (operational or maintenance) cost. Thus, the running costs and the resale price of the machine in successive years are as follows:

Year:	1	2	3	4	5
Resale Value (Rs.)	42, 000	30, 000	20, 400	14, 400	9, 650
Running Cost (Rs.)	18, 000	20, 270	22, 880	26, 700	31, 800

The calculations of average running cost per year during the life of the machine are shown in table 1.

Year of Service n	Running Cost (Rs.) R(n)	Cumulative Running Cost (Rs.) Σ R(n)	Resale Value (Rs.) S	Depreciation Cost (Rs.) C - S	Total Cost (Rs.) TC	Average Cost (Rs.) ATC_n
1	2	3	4	5 = 60, 000	6 = 3 + 5	7 = 6/1
1	18, 000	18, 000	42, 000	18, 000	36, 000	36, 000.00
2	20, 270	38, 270	30, 000	30, 000	68, 270	34, 135.00
3	22, 880	61, 150	20, 400	39, 600	1, 00, 750	33, 583.00
4	26, 700	87, 850	14, 400	45, 600	1, 33, 450	33, 362.00
5	31, 800	1, 19, 650	9, 650	50, 350	1, 70, 000	34, 000.00

The calculations in table 1 reveal that the average cost is lowest during the fourth year. Hence, the machine should be replaced after every fourth year, otherwise the average cost per year for running the machine would start increasing.

3.4. REPLACEMENT OF ITEMS THAT FAIL COMPLETELY

It is somehow difficult to predict that a particular equipment will fail at a particular time. This uncertainty can be avoided by deriving the probability distribution of failures. Here, it is assumed that the failures occur only at the end of the period, say t. Thus, the objective becomes to find the value of t which minimises the total cost involved for the replacement.

Mortality Tables: These tables are used to derive the probability distribution of life span of an equipment in question. Let

M (f) = Number of survivors at any time t

M (t - 1) = Number of survivors at any time t - 1

N = Initial number of equipments

Then the probability of failure during time period t is given by

$P(t) = \frac{M(t-1) - M(t)}{N}$

The probability that an equipment has survived to an age (t - 1). and will fail during the interval (t - 1) to t can be defined as the conditional probability of failure. It is given by

$P_c(t) = \frac{M(t-1) - M(t)}{M(t-1)}$

The probability of survival to an age t is given by

$P_s(t) = \frac{M(t)}{N}$

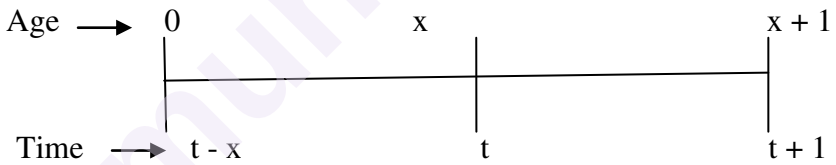
Mortality Theorem 1: A large population is subject to a given mortality law for a very long period of time. All deaths are immediately replaced by births and there are no other entries or exits. Show that the age distribution ultimately becomes stable and that the number of deaths per unit time becomes constant and is equal to the size of the total population divided by the mean age at death.

Proof: Without any loss of generality, it is assumed that death (or failure) occurs just before the age of (k + 1) years, where k is an integer. That is, life span of an item lies between t = 0 and t = k. Let us define f(t) = number of births (replacements) at time t, and

p(x) = probability of death (failure) just before the age x + 1, i.e. failure at time x.

and $\sum_{x=0}^{k} p(x) = 1$

If f(t - x) represents the number of births at time t - x, t = k, k +1, k + 2,....then the age of newly borns attain the age x at time t illustrated in the figure below:



Hence the expected number of deaths of such newly borns before reaching the time t + 1 (i.e. at time t) will be

Expected number of death = $\sum_{x=0}^{k} p(x) f(t - x)$, t = k, k + 1, ...

Since all deaths (failures) at time t are replaced immediately by births (replacements) at time t + 1, expected number of births are:

$$f(t + 1) = \sum_{x=0}^{k} p(x) f(t - x), \quad t = k, k + 1, \ldots\ldots \tag{1}$$

The solution to the difference Eq. (9) in t can be obtained by putting the value $f(t) = A\alpha^{t}$, where A is some constant. Then Eq. (9) becomes

$$A\alpha^{t+1} = A \sum_{x=0}^{k} p(x)\, \alpha^{t-x} \tag{2}$$

Dividing both sides of Eq. (10) by $A\alpha^{t-k}$, we get

$$\alpha^{k+1} = \sum_{x=0}^{k} p(x)\, \alpha^{k-x} = \alpha^{k} \sum_{x=0}^{k} p(x)\, \alpha^{-x}$$

$$= \alpha^{k} \{p(0) + p(1)\alpha^{-1} + p(2)\alpha^{-2} + \ldots\ldots\}$$

$$\text{or } \alpha^{k+1} - \{p(0)\alpha^{k} + p(1)\alpha^{k-1} + p(2)\alpha^{k-2}\} = 0 \tag{3}$$

Equation (3) is of degree (k + 1) and will, therefore, have exactly (k + 1) roots. Let us denote the roots of Eq. (3) by $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_k$.

For $\alpha = 1$, the L. H. S of equation (3), becomes

$$\text{L.H.S} = 1 - \{p(0) + p(1) + p(2) + \ldots + p(w)\}$$
$$= 1 - \sum_{x=0}^{w} p(x) = 0 \text{ R. H.S}$$

Hence, one root of Eq. (3) is $\alpha = 1$. Let us denote this root by α_0. The general solution of Eq. (3) will then be of the form

$$f(t) = A_0\, \alpha_0^t + A_1\, \alpha_1^t + \ldots + A_k\, \alpha_k^t$$
$$= A_0 + A_1\, \alpha_1^t + \ldots + A_k\, \alpha_k^t$$

where $A_0, A_1, A_2, \ldots A_k$ are constant whose values are to be calculated. (4)

Since one of the roots of Eq. (3), $\alpha_0 = 1$ is positive, according to the Descrae's sign rule all other roots $\alpha_1, \alpha_2, \ldots, \alpha_k$ will be negative and their absolute value is less than unity, i.e. $|\alpha_i| < 1$, i = 1, 2, 3, ..., k. It follows that the value of these roots tends to be zero as $t \rightarrow \infty$. With the result that Eq. (4) becomes $f(t) = A_0$. This indicates that the number of deaths (as well as births) becomes constant at any time.

Now the problem is to determine the value of the constant A_0. For this we can proceed as follows:

Let us define

$$g(x) = \text{Probability of survival for more than x years.}$$
$$\text{or } g(x) = 1 - \text{Prob (survivor will die before attaining the age x)}$$
$$= 1 - \{p(0) + p(1) + \ldots p(x-1)\}$$

Obviously, it can be assumed that g(0) = 1.

Since the number of births as well as deaths has become constant and equal to A_0, expected number of survivors of age x is given by $A_0 \,.\, g(x)$.

As deaths are immediately replaced by births, size N of the population remains constant. That is,

$$N = A_0 \sum_{x=0}^{k} g(x)\ \alpha_1, \alpha_2, \ldots, \alpha_k$$

Or

$$A_0 = \frac{N}{\sum_{x=0}^{k} g(x)} \quad \ldots\ldots\ldots\ldots \quad (4)$$

The denominator in Eq. (4) represents the average age at death. This can also be proved as follows:

From finite differences, we know that

$$\Delta(x) = (x + 1) - x = 1$$
$$\sum_{x=a}^{b} f(x)\, \Delta h(x) = f(b + 1)\, h(b + 1) - f(a)\, h(a) - \sum_{x=a}^{b} (x + 1)\, \Delta g(x)$$
$$= g(k + 1)\,(k + 1) - g(0)\,.\,0 - \sum_{x=0}^{k} (x + 1)\, \Delta g(x)$$
$$= g(k + 1)\,(k + 1) - \sum_{x=0}^{k} (x + 1)\, \Delta g(x) \ \ldots\ldots\ldots\ldots \quad (5)$$
$$\text{But } g(k + 1) = 1 - \{p(0) + p(1) + p(2) + \ldots + p(k)\} = 0$$

since no one can survive for more than (k + 1) years of age and

$$\Delta g(x) = g(x + 1) - g(x)$$
$$= \{1 - p(0) - p(1) - \ldots - p(x)\} - \{1 - p(0) - p(1) \ldots p(x - 1)\} = -\, p(x)$$

Substituting the value of g(k + 1) and $\Delta g(x)$ in Eq. (5), we get

$$\sum_{x=0}^{k} g(x) = \sum_{x=0}^{k} (x + 1) p(x) = \text{Mean age at death}$$

Hence from Eq. (4), we get

$$A_0 = \frac{N}{Average\ age\ at\ death}.$$

STAFFING PROBLEM

The principles of replacement may be applied to formulate some useful recruitment and promotion policies for the staff working in the organisation. For this, we assume that life distribution for the services of staff in the organisation is already known.

Example 1:

An airline requires 200 assistant hostesses, 300 hostesses and 50 supervisors. Women are recruited at the age of 21, and if still in service retire at 60. Given the following life table, determine

(a) How many women should be recruited in each year?

(b) At what age should promotion take place?

Airline Hostesses' Life Record

Age	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**
No. in Servie	**1000**	**600**	**480**	**384**	**307**	**261**	**228**	**206**
Age	29	30	31	32	33	34	35	36
No. in Servie	190	181	173	167	161	155	150	146
Age	37	38	39	40	41	42	43	44
No. in Servie	141	136	131	125	119	113	106	99
Age	45	46	47	48	49	50	51	52
No. in Servie	93	87	80	73	66	59	53	46
Age	53	54	55	56	57	58	59	---
No. in Servie	39	33	27	22	18	14	11	

Solution: If 1000 women had been recruited each year for the past 39 years, then the total number of them recruited at the age of 21 and those serving upto the age of 59 is 6, 480. Total number of women recruited in the airline are: 200 + 300 + 50 = 550.

(a) Number of women to be recruited every year in order to maintain a strength of 55 hostesses 550 x (1000/6840) = 85 approx.

(b) If the assistant hostesses are promoted at the age of x, then up to age (x - 1), 200 assistant hostesses will be required. Among 550 women, 200 are assistant hostesses. Therefore, out of a strength of 1, 000 there will be: 200 x (1000/550) = 364 assistant hostesses. But, from the life table given in the question, this number is available up to the age of 24 years. Thus, the promotion of assistant hostesses is due in the 25th year.

Since out of 550 recruitments only 300 hostesses are needed, if 1,000 girls are recruited, then only 1000 x (300/550) = 545 approx. will be hostesses. Hence, total number of hostesses and assistant hostesses in a recruitment will be: 545 + 364 = 909. This means, only 1000 - 909 = 91 supervisors are required. But from life table this number is available up to the age of 46 years. Thus promotion of hostesses to supervisors will be due in 47th year.

3.5. EXERCISE

1. In the theory of replacement models construct an equation for the cost of maintaining a system as a function of the control variable t (the number of periods between group replacements).

2. State some of the simple replacement policies and give the average cost functions for the same explaining your notations.

3. The cost of maintenance of a machine is given as a function that the average annual cost will be minimised by replacing the machine when the average cost to date becomes equal to the current maintenance cost.

4. Discuss the importance of replacement models.

5. Explain how the theory of replacement is used in the following problems.
(i) Replacement of items whose maintenance cost varies with time.
(ii) Replacement of items that fail completely.

6. Find the cost per period of individual replacement of an installation of 300 lighting bulbs, given the following
(a) cost of replacing individual bulb is Rs. 3
(b) Conditional probability of failure is given below:

Week Number	:	0	1	2	3	4
Conditional Probability of Failure	:	0	1/10	1/3	2/3	1

7. The following mortality rates have been observed for a special type of light bulbs.

Month	:	1	2	3	4	5
% failing at the end of the month	:	10	25	50	80	100

In an industrial unit there are 1,000 special type of bulbs in use, and it costs Rs.10 to replace an individual bulb which has burnt out. If all bulbs were replaced simultaneously it would cost Rs. per bulb. It is proposed to replace all bulbs at fixed intervals, whether or not they have burnt out, and to continue replacing burnt out bulbs as they fail. At what intervals of time should the manager replace all the bulbs?

8. An airline, whose staff are subject to the same survival rates as in the previous problem, currently has a staff whose ages are distributed in the following table. It is estimated that for the next two years' staff requirements will increase by 10% per year. If women are to be recruited at the age of 21, how many should be recruited for the next year and at what age will promotions take place? How many should be recruited for the following year and at what age will promotions take place?

munotes.in

Assistant Age Number	**21 90**	**22 50**	**23 30**	**24 20**	**25 10**		**(Total 200)**		
Hostesses Age Number	26 40	27 35	28 35	29 30	30 28	31 26	32 20	33 18	34 16
Age Number	35 12	36 10	37 8	38 ---	39 8	40 8	41 6 (Total 300)		
Supervisors Age Number Age Number	42 5 51 ---	43 4 52 4	44 5 53 3	45 3 54 5	46 3 55 ---	47 3 56 3	48 6 57 2	49 2 58 ---	50 --- 59 2 (Total 50)

munotes.in

munotes.in

CHAPTER 4

INVENTORY MODELS

UNIT STRUCTURE

4.1 Introduction
4.2 Importance of Inventory Control
4.3 Inventory Control Decisions
4.4 Inventory Control with Deterministic Models
4.5 Selective Inventory Control techniques- ABC Analysis
4.6 VED Analysis
4.7 FSN Analysis
4.8 XYZ Analysis
4.9 Exercise

munotes.in

4.1. INTRODUCTION

Inventory is one of the most expensive and important assets of many companies, representing as much as 50% of total invested capital. Managers have long recognised that good inventory control is crucial. On one hand, a firm can try to reduce costs by reducing on-hand inventory levels. On the other hand, customers become dissatisfied when frequent inventory outages, called stock outs occur. Thus, companies must make the balance between low and high inventory levels. As you would expect, cost minimisation is the major factor in obtaining this delicate balance.

Inventory is any stored resource that is used to satisfy a current or future need. Raw materials, work-in-progress, and finished goods are examples of inventory. Inventory levels for finished goods, such as clothes dryers, are a direct function of market demand. By using this demand information, it is possible to determine how much raw materials (eg., sheet metal, paint, and electric motors in the case of clothes dryers) and work-in-processes are needed to produce the finished product.

Every organisation has some type of inventory planning and control system. A bank has methods to control its inventory of cash. A hospital has methods to control blood supplies and other important items. State and federal governments, schools, and virtually every manufacturing and production organisation are concerned with inventory planning and control. Studying how organisations control their inventory is equivalent to studying how they achieve their objectives by supplying goods and services to their customers. Inventory is the common thread that ties all the functions and departments of the organisation together.

Figure 4.1 illustrates the basic components of inventory planning and control system. The planning phase involves primarily what inventory is to be stocked and how it is to be acquired (whether it is to be manufactured or purchased). This information is then used in forecasting the demand for the inventory and in controlling inventory levels. The feedback loop in figure 4.1 provides a way of revising the plan and forecast based on experiences and observation.

Through inventory planning, an organisation determines what goods and/or services are to be produced. In cases of physical products, the organisation must also determine whether to produce these goods or to purchase them from another manufacturer. When this has been determined, the next step is to forecast the demand. Many mathematical techniques can be used in forecasting demand for a particular product. The emphasis in this chapter is on inventory control - that is, how to maintain adequate inventory levels within an organisation to support a production or procurement plan that will satisfy the forecasted demand.

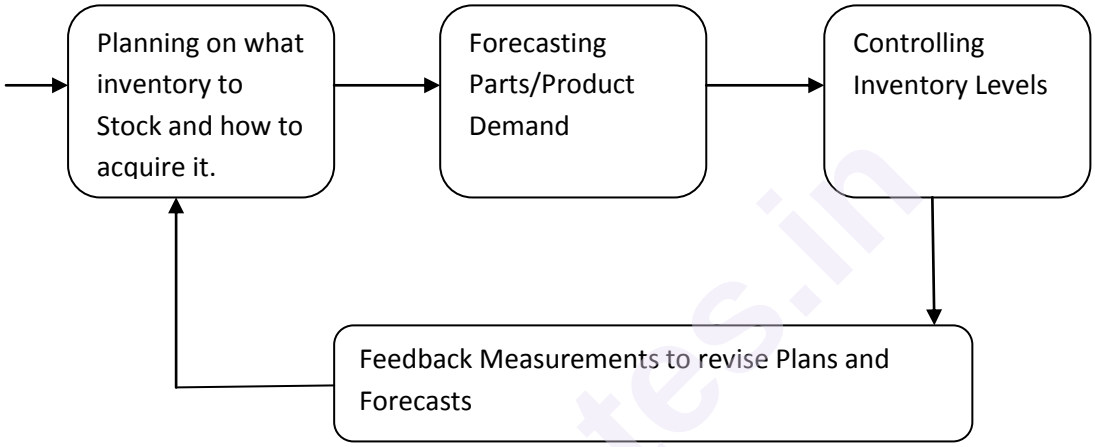
In this chapter, we discuss several different inventory control models that are commonly used in practice. For each model, we provide examples of how they are analysed. Although we show the equations needed to compute the relevant parameters for each model, we use EXCEL worksheets to actually calculate these values.

4.2. IMPORTANCE OF INVENTORY CONTROL

Inventory control serves several important functions and adds a great deal of flexibility to the operation of a firm. Five main uses of inventory are as follows:

1. Decoupling function
2. Storing resources
3. Irregular supply and demand
4. Quantity discounts
5. Avoiding stockouts and shortages.

Figure 4.1 Inventory Planning and Control



1. Decoupling Function

One of the major functions of the inventory is to decouple manufacturing processes within the organisation. If a company did not store inventory, there could be many delays and inefficiencies. For example, when one manufacturing activity has to be completed before a second activity can be started, it could stop the entire process. However, stored inventory between processes could act as a buffer.

2. Storing Resources

Agricultural and seafood products often have definite seasons over which they can be harvested or caught, but the demand for these products is somewhat constant during the year. In these and similar cases, inventory could be used to store these resources.

In manufacturing process, raw materials can be stored by themselves, as work-in-process, or as finished products. Thus, if your company makes lawn movers, you may obtain lawn mower tyres from another manufacturer. If you have 400 finished lawn movers and 300 tyres in inventory, you actually have 1, 900 tyres stored in inventory. 300 tyres are stored by themselves, and 1, 600 (= 4 tyre per lawn mower X 400 lawn movers) tyres are stored on the finished lawn mowers. In the same sense, labour can be stored in inventory. If you have 500 subassemblies and it takes 50 hours of labour to produce each assembly, you actually have 25, 000 labour hours stored in inventory in the subassemblies. In general, any resource, physical or otherwise can be stored in inventory.

3. Irregular supply and demand

When the supply and demand for the inventory item is irregular, storing certain amount as inventory can be important. If the greatest demand for Diet-Delight beverage is during the summer, the Diet-Delight company will have to make sure that there is enough supply to meet this irregular demand. This might require that the company produce more of the soft drink in the winter than is actually needed in order to meet the winter demand. The inventory levels of Diet-Delight will gradually build up over the winter, but this inventory will be needed in summer. The same is true for irregular supplies.

4. Quantity Discounts

Another use of inventory is to take advantage of quantity discounts. Many suppliers offer discounts for larger orders. For example, an electric jigsaw might normally cost $10 per unit. If you order 300 or more saws at one time, your supplier may lower the cost to $8.75. Purchasing in larger quantities can substantially reduce the cost of products. There are, however, some disadvantages of buying in larger quantities. You will have higher storage costs and higher costs due to spoilage, damaged stock, theft, insurance, and so on. Furthermore, if you invest in more inventory, you will have less cash to invest elsewhere.

5. Avoiding stock outs and shortages

Another important function in inventory is to avoid stockouts and shortages. If a company is repeatedly out of stock, customers are likely to go elsewhere to satisfy their needs. Lost goodwill can be an expensive price to pay for not having the right item at the right time.

4.3. INVENTORY CONTROL DECISIONS

Even though there are literally millions of different types of products manufactured in our society, there are only two fundamental decisions that you have to make when controlling inventory:

1. How much to order
2. When to order

Table 4.1 Inventory Cost Factors

Operating Cost Factors	Carrying Cost Factors
Developing and sending purchase orders Processing and inspecting incoming inventory Bill Paying Inventory Inquiries Utilities, phone bills, and so on for the purchasing department Salaries and wages for purchasing department employees Supplies such as forms and paper for the purchasing department	Cost of Capital Taxes Insurance Spoilage Theft Obsolescence Salaries and wages for warehouse employees Utilities and building costs for the warehouse Supplies such as forms and papers for the warehouse

The purpose of all inventory models is to determine how much to order and when to order. As you know, inventory fulfils many important functions in an organisation. But as the inventory levels go up to provide these functions, the cost of holding and storing inventories also increases. Thus, we must reach a fine balance in establishing inventory levels. A major objective in controlling inventory is to minimise total inventory costs. Some of the most significant inventory costs are as follows:

1. Cost of the items
2. Cost of ordering
3. Cost of carrying or holding inventory
4. Cost of stock outs
5. Cost of safety stock, the additional inventory that may be held to help avoid stockouts.

The inventory models discussed in the first part of this chapter assume that demand and the time it takes to receive an order are known and constant, and that no quantity discounts are given. When this is the case, the most significant costs are the cost of placing an order and the cost of holding inventory items over a period of time. Table 4.1 provides a list of important factors that make up these costs.

ECONOMIC ORDER QUANTITY: DETERMINING HOW MUCH TO ORDER

The Economic Order Quantity (EOQ) model is one of the oldest and most commonly known inventory control techniques. Research on its use dates back to a 1915 publication by Ford W. Harris. This model is still used by a large number of organisations today. This technique is relatively easy to use, but it makes a number of assumptions. Some of the more important assumptions are as follows:

ASSUMPTIONS OF EOQ

1. Demand is known and constant.
2. The lead time - that is, the time between the placement of the order and the receipt of the order - is known and constant.
3. The receipt of inventory is instantaneous. In other words, the inventory from an order arrives in one batch, at one point in time.
4. Quantity discounts are not possible.
5. The only variable costs are the cost of placing an order, ordering cost, and the cost of holding or storing inventory over time, carrying or holding cost.
6. If orders are placed at the right time, stock outs and shortages can be completely avoided.

With these assumptions, inventory usage has a sawtooth shape, as in Figure 4.2. Here, Q represents the amount that is ordered. If this amount is 500 units, all 500 units are arrived at one time when an order is received. Thus, the inventory level jumps from 0 to 500 units. In general, the inventory level increases from 0 to Q units when an order arrives.

Because demand is constant over time, inventory drops at a uniform rate over time. Another order is placed such that when the inventory level reaches 0, the new order is received and the inventory level again jumps to Q units, represented by the vertical lines. This process continues indefinitely over time.

OPTIMISING INVENTORY AT PROCTER & GAMBLE

Procter & Gamble (P&G) is a world leader in consumer products with annual sales of over $76 billion. Managing inventory in such a large and complex organisation requires making effective use of the right people, organisational structure, and tools. P&G's logistics planning personnel coordinate material flow, capacity, inventory, and logistics for the firm's extensive supply chain network, which comprises P&G-owned manufacturing facilities, 300 contract manufacturers, and 6, 900 unique product-category market combinations. Each supply chain requires effective management based on the latest available information, communication and planning tools to handle complex challenges and trade-offs on issues such as production batch sizes, order policies, replenishment timing, new product introductions and assortment management.

Though the effective use of inventory optimisation models, P&G has reduced its total inventory investment significantly. Spreadsheet based inventory models that locally optimise different portions of the supply chain drive nearly 60 % of P&G's business. For more complex supply chain networks (which drive about 30% of P&G's business), advanced multi-stage models yield additional average inventory reductions of 7%. P&G estimates that the use of these tools was instrumented in driving $1.5 billion in cash savings in 2009, while maintaining or increasing service levels.

Source: Based on I. Farasyn et. al. "Inventory Optimisation at P&G: Achieving real benefits through User Adoption of Inventory Tools," Interfaces 41, 1(January-February 2011): 66-78

munotes.in

ORDERING AND INVENTORY COSTS

The objective of most inventory models is to minimise the total cost. With the assumptions just given, the significant costs are the ordering costs and inventory carrying cost. All other costs, such as the cost of the inventory itself, are constant. Thus, if we minimise the sum of the ordering and carrying costs, we also minimise the total cost.

To help visualise this, Figure 4.3 graphs total cost as a function of the order quantity, Q. As the value of Q increases, the total number of orders placed per year decreases. Hence, the total ordering cost decreases. However, as the value of Q increases, the carrying cost increases because the firm has to maintain larger average inventories.

The optimal order size, Q^*, is the quantity that minimises the total cost. Note in figure 12.3 that Q^* occurs at the point where the ordering cost curve and the carrying cost curve intersect. This is not by chance. With this particular type of cost function, the optimal quantity always occurs at a point where the ordering cost is equal to the carrying cost.

Figure 12.2 Inventory Usage Over time

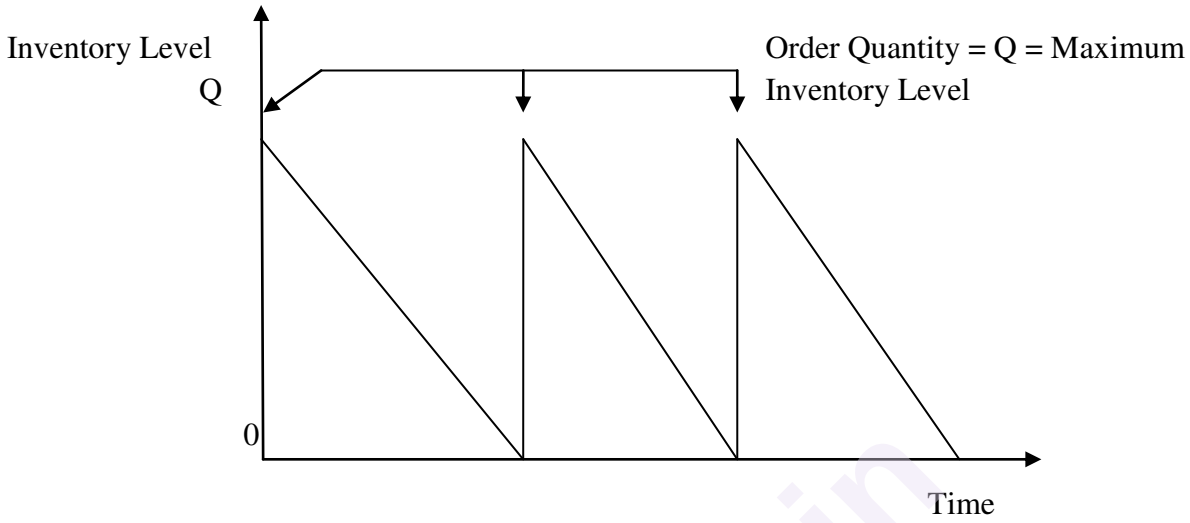
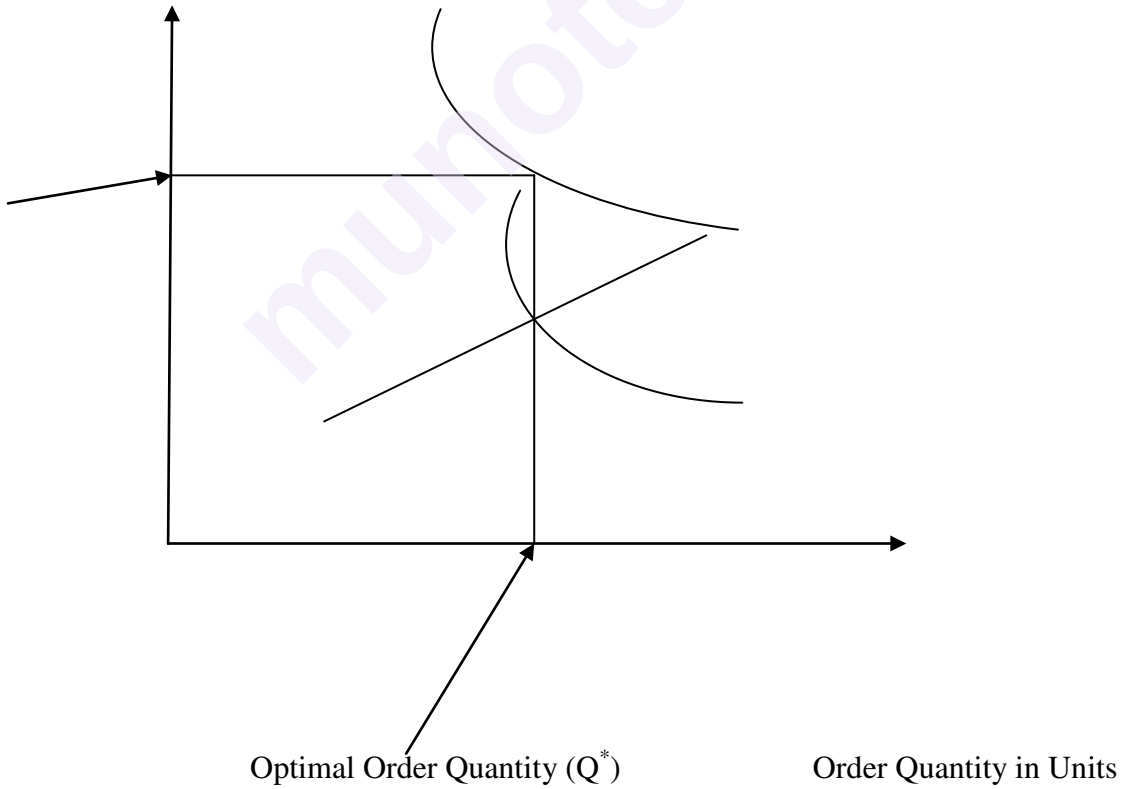


Figure 4.3 Total Cost as a Function of Order Quantity



Now that we have a better understanding of inventory costs, let us see how we can determine the value of Q^* that minimises the total cost. In determining the annual carrying cost, it is convenient to use the average inventory. Referring to figure 4.2, we see that the on-hand

inventory ranges from a high of Q units to a low of zero units, with a uniform rate of decrease between these levels. Thus, the average inventory can be calculated as the average of the minimum and maximum inventory levels. That is,

$$\text{Average Inventory} = (0 + Q)/2 = Q/2 \text{ ---------- } (1)$$

We multiply this average inventory by a factor called the annual inventory carrying cost per unit to determine the annual inventory cost.

FINDING THE ECONOMIC ORDER QUANTITY

We pointed out that the optimal order quantity, Q^*, is the point that minimises the total cost, where total cost is the sum of ordering cost and carrying cost. We also indicated graphically that the optimal order quantity was at the point where the ordering cost was equal to the carrying cost. Let us now define the following parameters:

Q^* = Optimal Order Quantity (i.e. EOQ)

D = Annual Demand, in units, for the inventory item

C_o = Ordering cost per order

C_h = Carrying or holding cost per unit per year

P = Purchasing cost per unit of the inventory item

The unit carrying cost, C_h, is usually expressed in one of two ways, as follows:

1. As a fixed cost, for example, C_h is \$0.50 per unit per year.
2. As a percentage (typically denoted by l) of the item's unit purchase cost or price. For example, C_h is 20% of the item's unit cost. In general

$$C_h = I \times P \text{ ---------- } (2)$$

For a given order quantity, Q, the ordering, holding and total costs can be computed using the following formulae:

$$\text{Total Ordering Cost} = (D/Q) \times C_o \text{ ---------- } (3)$$

$$\text{Total Carrying Cost} = (Q/2) \times C_h \text{ ---------- } (4)$$

$$\text{Total Cost} = \text{Total Ordering Cost} + \text{Total Carrying Cost} + \text{Purchase Cost}$$

$$= (D/Q) \times C_o + (Q/2) \times C_h + P \times D \text{ ---------- } (5)$$

Observe that the total purchase cost (i.e. P x D) does not depend on the value of Q. This is so because regardless of how many orders we place each year, or how many units we order each time, we will still incur the same annual total purchase cost.

The presence of Q in the denominator for the first term makes equation (5) a non-linear equation with respect to Q. Nevertheless, because the total ordering cost is equal to the total carrying cost at the optimal value of Q, we can set the terms in equation (3) and (4) equal to each other and calculate EOQ as

$$Q^* = \sqrt{2D\, C_o / C_h} \text{ ---------- } (6)$$

4.4. INVENTORY CONTROL WITH DETERMINISTIC MODELS

Examples 1

1. The production department for a company requires 3,600 kg of raw material for manufacturing a particular item per year. It has been estimated that the cost of placing an order is Rs. 36 and the cost of

carrying inventory is 25% of the investment in the inventories. The price is Rs. 10/kg. The purchase manager wishes to determine an ordering policy for raw material.

Solution

From the data of the problem we know that

D = 3600 kg per year; C_o = Rs. 36 per order; C_h = 25% of the investment in inventories = 10 X 0.25 = Rs. 2.50 per kg per year.

a. The optimal lot size is given by

$Q^* = \sqrt{2D}\ C_o / C_h = \sqrt{2}\ X\ 3600\ X\ 36 / 250 = 321.99$ kg per order.

b. Optimal order cycle time

$t^* = \frac{Q*}{D} = 321.99/3600 = 0.894$ year

c. The minimum yearly variable inventory cost

$TVC = \sqrt{2D}\ C_oC_h = \sqrt{2}X\ 3600\ X\ 36\ X\ 2.5$ = Rs. 804.98 per year

d. The minimum yearly total inventory cost

$TC^* = TVC^* + DC$ = Rs. 804-98 + (3600 kg) (Rs. 10/kg) = Rs. 36, 804.98 per year.

2. A company operating 50 weeks in a year is concerned about its stocks of copper cable. This costs Rs. 240/- a meter and there is a demand for 8, 000 meters a week. Each replenishment costs Rs. 1,050 for administration and Rs. 1, 650 for delivery, while holding costs are estimated at 25% of value held a year. Assuming no shortages are allowed, what is the optimal inventory policy for the company?

How would this analysis differ if the company wanted to maximise profit rather than minimise cost? What is the gross profit if the company sell cable for Rs. 360 a meter?

Solution

From the data of the problem, we have

Demand rate (D) = 8, 000 x 50 = 4, 00, 000 metres a year

Purchase Cost (C) = Rs. 240 a unit; Ordering cost (C_o) = 1, 050 + 1, 650 = Rs. 2, 700

Holding Cost = 0.25 x 240 = Rs. 60 a meter a year

a. Optimal order quantity $Q^* = \sqrt{2D}\ C_o / C_h = \sqrt{\frac{2\ X\ 40{,}000\ X\ 2{,}700}{60}}$ = 6, 000 meters

b. Total variable inventory cost, $TVC = Q^*.\ C_h$ = 6, 000 x 60 = Rs. 3, 60, 000 a year

c. Total inventory cost, TC = D.C + TVC = 4, 00, 000 x 240 + 3, 60, 000 = Rs. 9, 63, 60, 000

With a turnover in excess of 96 million rupees a year inventory costs are only Rs. 3, 60, 000 or 0.36%. This figure is usually low but any well run organisation should try to make all the waiving it can, however, small.

If the company wanted to maximise profit rather than minimise cost, the analysis used would remain exactly the same. This can be demonstrated by defining selling price (SP) per unit so that gross profit per unit becomes,

Profit = Revenue - Cost

$$= D \times SP - \left(DC + \frac{D}{Q} C_o + \frac{Q}{2} C_h \right)$$

When this equation is solved to maximise profit with respect to Q as discussed earlier, we get the same result by applying usual method.

If company sells the cable for Rs. 360 a meter, its revenue is Rs. 360 x 4, 00, 000 = Rs. 14, 40, 00, 000 a year. Direct cost of Rs. 9, 63, 60, 000 is subtracted from this to get a gross profit of Rs. 4, 76, 40, 000 a year.

3. A chemical company is considering the optimal batch size for reorder of concentrated sulphuric acid. The management accountant has supplied the following information:

(1) The purchase price of H_2SO_4 is Rs. 150 per gallon.

(2) The clerical and data processing costs are Rs. 500 per order.

All the transport is done by rail. A charge of Rs. 2, 000 is made each time the special line to the factory is opened. A charge of Rs. 20 gallons is also made. The company uses 40, 000 gallons per year. Maintenance costs of stock are Rs. 400 per gallon per year.

Each gallon requires 0.5 sq ft of storage space. If warehouse space is not used, it can be rented out to another company at Rs. 200 per sq ft per annum. Available warehouse space is 1,000 sq ft, the overhead costs being Rs. 5, 000 p.a. Assume that all free warehouse space can be rented out.

(a) Calculate the economic reorder size

(b) Calculate the minimum total annual cost of holding and reordering stock.

Solution

Based on the data of the problem, the relevant cost components which will vary over the time period due to change in lot size quantity (Q) and which will remain fixed is summarised as follows:

	Relevant Cost for EOQ Calculation	**Irrelevant Cost for EOQ Calculation**
Ordering Cost	• Clerical and data processing, Rs. 500; • Rail transport, Rs. 2000	• Rail transport, Rs. 20 per gallon because a fixed money of Rs. 40, 000 x 20 = Rs. 8, 00, 000 will incur irrespective of size of Q.
Carrying cost	• Maintenance cost, Rs. 200; • Rented cost, Rs. 200/2 = Rs. 100	• Overhead cost, Rs. 5, 000

Thus the relevant costs needed for calculating EOQ are: C_o = Rs. 2, 500; C_h = Rs. 300.

Q^* (EOQ) = $\sqrt{\frac{2DCo}{Ch}} = \sqrt{\frac{2 \, x \, 40{,}000 \, x \, 2{,}500}{300}}$ = Rs. 817 (approx.) gallons.

Hence, the minimum total annual costs are:

Ordering: $\frac{D}{Q*} C_o = \frac{40{,}000}{817}$ x 2, 500 = Rs. 1, 22, 399.02

Carrying: $\frac{Q*}{2} C_h = \frac{817}{2}$ x 300 = Rs. 1, 22, 250.00

Total = Rs. 2, 44, 949.02

Rail transport: 40, 000 x 2 = Rs. 8, 00, 000

Storage Overhead cost: = Rs. 5, 000

Purchase Cost: 40, 000 x 150 = Rs. 60, 00, 000

Total = Rs. 70, 49, 949.02

4. A contractor has to supply 10, 000 bearings per day to an automobile manufacturer. He finds that when he starts production run, he can produce 25, 000 bearings per day. The cost of holding a bearing in stock for a year is Rs. 2 and the set-up cost of a production run is Rs. 180. How frequently should production run be made?

Solution

From the data of the problem in usual notation, we have

C_o = Rs. 1, 800 per production run; C_h = Rs. 2 per year

p = Rs. 25, 000 bearings per day d = 10, 000 bearings per day

D = 10, 000 x 300 = 30, 00, 000 units/year (assuming 300 working days in the year).

a. Economic batch quantity for each production run is given by

$Q^* = \sqrt{\frac{2DCo}{Ch}}\left(\frac{p}{p-d}\right) = \sqrt{\frac{2\ x\ 30{,}00{,}000\ x\ 1{,}800}{2}}\left(\frac{25{,}000}{25{,}000-10{,}000}\right)$ = 1, 04, 446 bearings

b. Frequency of production cycles

$t^* = \frac{Q*}{d}$ = 1, 04, 446/10, 000 = 10.44 days.

5. A commodity is to be supplied at a constant rate of 25 units per day. A penalty cost is being charged at the rate of Rs. 10 per unit per day late for missing the scheduled delivery date. The cost of carrying the commodity in inventory is Rs. 16 per unit per month. The production process is such that each month (30 days) a batch of items is started and are available for delivery any time after the end of the month. Find the optimal level of inventory at the beginning of each month.

Solution:

From the data in usual notations, we have

D = 25 units / day; C_h = 16/30 = 0.53 per unit per day; C_o = Rs. 10 per unit per day;

t = 30 days

Thus optimal inventory level is given by

$M^* = \left[\frac{10}{0.53+10}\right]$ (25) (30) = 712 units.

4.5. SELECTIVE INVENTORY CONTROL TECHNIQUES

In practice when a firm maintained large number of items in its inventory, obviously all items cannot, and need not be controlled (i.e. keeping record of time interval between successive reviews of demand; order frequencies; expected demand rate; order quantities etc.) with equal attention. All items are not of equal importance to the firm in such terms as sales, profits, availability etc. One way of exercising proper degree of control overall and various types of items held in stocks is to classify them into groups (or classes) on the basis of the degree of control or intensity of management effort that they require.

By selectively applying inventory control policies to these different groups, inventory objectives can be achieved with lower inventory levels than with a single policy applied to all items. These techniques are also known as selective multi-item inventory control techniques.

In this section, we shall consider certain group classifications such as: ABC, VED, FSN, HML, XYZ etc.

Classification	Basis of Classification	Purpose
ABC {Always, Better, Control}	Value of consumption	To control raw material, components and work-in-progress inventories in the normal course of business.
HML {High, Medium, Low}	Unit price of the material	Mainly to control purchases

XYZ	Value of items in storage	To review the inventories and their uses at scheduled intervals.
VED {Vital, Essential, Desirable}	Criticality of the component	To determine the stocking levels of spare parts.
FSN {Fast, Slow, Non-moving}	Consumption pattern of the component	To control obsolescence

ABC ANALYSIS

The ABC analysis consists of separating the inventory items into three groups: A, B and C according to their annual cost volume consumption (unit cost x annual consumption). Although the break points between these groups vary according to individual business conditions, a common break down might be as follows:

Category or Group	**Percentage of the item**	**% of the Total Annual Value of the inventories (Rs.)**
A	10 - 20	70 - 85
B	20 - 30	10 - 25
C	60 - 70	5 - 15

This type of classification is also known as the principle of law of Vital Few and Trivial Many. The ABC analysis facilitates analysis of yearly consumption value of items in the store to identify the vital few items which are generally referred to as A category items. Generally, these items accounting for about 70% of the total money value of consumption. Items accounting for about 25% of the total money value of consumption are called B category items and the remaining ones accounting for about 5% consumption value as C category items.

Carrying on the ABC analysis of the store items helps identifying the few items that are vital from financial point of view and require careful watch, scrutiny and follow-up. The application of ABC analysis extends overall of the aspects of materials management like purchasing, inventory control, value analysis etc.

After the items are so classified, the inventory control policies are made on the basis of the classification 'A' category items require special managerial attention, therefore, fixed-interval inventory control system might be used for these items. 'C' category items can be managed in a little casual manner. For these items, a fixed-order quantity system might be used. The order quantities can be relatively large without incurring excessive costs. A large reserve stock can also be maintained. 'B' items are not so costly as to require special managerial attention, but these are not so cheap as to ignore overstocking, therefore (s, S) inventory control system might be used for these items.

The procedure of ABC analysis is summarised in the following steps:

Step 1

Obtain data on the annual usage (or consumption) in units and unit cost of each inventory unit. Multiple the annual usage in units and the value of each item to get annual value for each of these items.

Annual Value = Unit Cost x Annual Consumption

Step 2

Arrange these inventory items in a decreasing order of their value computed in step 1.

Step 3

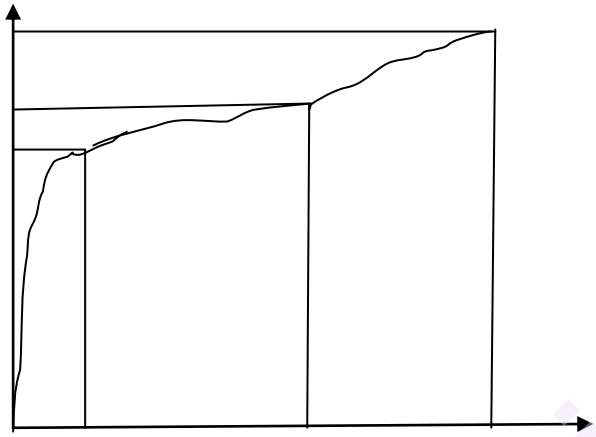
Express the annual value of each item as percentage of the total value of all items. Also compute the cumulative percentage of annual consumption rupees spent.

Step 4

Obtain the percentage value for each of the items. That is, if there are 50 items involved in classification, then each item would represent 100/50 = 2 percent of the total items. Also cumulate these percentage values.

Step 5

Draw a graph between cumulative percentage of items (on x-axis) and cumulative annual percentage of usage value (on y-axis), and mark cut-off points where the graph changes slope as shown in figure.



Example 1

A company produces a mix of high technology products for use in hospitals. The annual sales data are as follows:

Product Type	Number of Units	Unit Price (Rs.)	Product Type	Number of Units	Unit Price (Rs.)
1	1, 000	2.50	10	600	1.62
2	250	0.55	11	25	33.00
3	150	6.50	12	4	15.50
4	300	1.00	13	1, 000	5.00
5	100	1.50	14	2, 850	2.50
6	700	1.43	15	10	0.83
7	500	7.00	16	355	0.98
8	15	4.98	17	50	1.37
9	1, 000	0.75	18	393	1.85

For inventory control reasons, the company wants to classify these items into three groups A, B and C on the basis of annual sales value of each item. You please help the company.

Solution

The annual sales volume (in Rs.) for each product and the item ranking on the basis of this volume is shown in Table.

Product ranking as per sales volume

Product Type	**Number of Units**	**Unit Price (Rs.)**	**Annual Sales Volume (Rs.)**	**Ranking**
1	1000	2.50	2, 500.00	4
2	250	0.55	137.50	14
3	150	6.50	975.00	6
4	300	1.00	300.00	12
5	100	1.50	150.00	13
6	700	1.43	1001.00	5

7	500	7.00	3500.00	3
8	15	4.98	74.70	16
9	1000	0.75	750.00	9
10	600	1.62	972.00	7
11	25	33.00	825.00	8
12	4	15.50	77.50	15
13	1000	5.00	5000.00	2
14	2850	2.50	7125.00	1
15	10	0.83	8.30	18
16	355	0.98	347.90	11
17	40	1.37	54.80	17
18	393	1.85	727.05	10

The cumulative percentage of products and cumulative percentage of sales for each product is given in table for the purpose of ABC classification:

ABC Classification

Rank	Product	Product Cumulative % of products	Annual Sales Volume (Rs.)	Cumulative Annual Sales Volume (Rs.)	Cumulative % Product Class of Sales
1	14	5.56	7, 125.00	7,125.00	29.05 A product 11.11%
2	13	11.11	5, 000.00	12,125.00	49.43 products and 49.43 Rs.
3	7	16.67	3, 500.00	15,625.00	63.70 B products: 38.89%
4	1	22.22	2, 500.00	18,125.00	73.90 Products and 42.91 Rs.
5	6	27.78	1, 001.00	19,125.00	77.97
6	3	33.33	975.00	20,101.00	81.95
7	10	38.89	825.00	21,073.00	85.92
8	11	44.44	750.00	21,898.00	89.28
9	9	50.00	727.05	22,648.00	92.34
10	18	55.56	347.90	23,375.00	95.30 C Products 44.44%
11	16	61.11	300.00	23,722.00	96.72 products and 7.66 Rs.
12	4	66.67	150.00	24,022.00	97.94
13	5	72.22	137.50	24,172.95	98.56
14	2	77.78	27.50	24,310.45	99.12
15	12	83.33	24.70	24,387.95	99.43
16	8	88.89	54.80	24,462.65	99.74
17	17	94.44	8.30	24,517.45	99.96
18	15	100.00		24,525.75	100.00

The percentage of products and percentage of annual sales volume can also be plotted on the graph.

munotes.in

4.6. VED Analysis

This analysis helps in separating the inventory items into three groups according to their criticality, usually called V, E, and D items in that order, VED classification calls for classification of items as Vital, Essential and Desirable.
V items are considered vital for smooth running of the system and without these items the whole system becomes inoperative. Thus adequate stock is required all the time.

E items are considered essential to the efficient running of the system and non-availability of these items reduces the efficiency of the system.

D items neither stop the system nor reduce its efficiency, but availability of such items will lead to increase in efficiency and reduction of failure.

This classification is largely useful in controlling inventory of spare parts. It can also be used in case of such raw materials whose availability is rare.
ABC analysis and VED analysis can also be combined to control the stocking of spare parts based on the desired customer service level as shown in the table.

ABC Classification	**VED Classification**		
	V	**E**	**D**
A	Constant control, Regular follow-up Low stocks and ordering more frequently	Average stock; No risk of stock outs.	No stock
B	Average stocks No risk of stock outs	Average stock; some risk can be taken	Very low stocks; Some risk can be taken.
C	High Stocks Restricted orders; No risk	Average Stock; some risk can be taken.	Low stocks; some risk can be taken.

HML Analysis
Based on the unit price of items, the HML classification separates inventory items, as High price, Medium price and Low price. This analysis is helpful to control purchase of various items for inventory.

4.7. FSN Analysis

The consumption pattern or inventory items forms the basis for FSN analysis. Items are classified as Fast-moving, Slow-moving and Non-moving. Sometimes items are also classified as FSND: Fast-moving; Slow-moving; Normal-moving and Dead (or Non-moving).

This classification is based on the movement (or consumption pattern) and therefore helps to controlling obsolescence of various items by determining the distribution and handling patterns. Cut-off points of the three classes are usually in terms of number of issues in the previous few years.

4.8. XYZ Analysis

This classification is based on the closing value of items in storage. Items whose inventory values are high and moderate are classified as X-items and Y-items respectively, while items with low inventory value are termed as Z-items.
This analysis is usually undertaken once a year during the annual stock taking exercise. This helps in identifying the items which are being stocked extensively.
This classification can also be combined with ABC classification of items to control inventory of items as shown in the table below:

ABC Classification	**XYZ Classification**		
	X	**Y**	**Z**
A	Attempt to reduce stocks	Attempt to convert Z-items	Items are within control
B	Stock and consumption is reviewed more often	Items are within control	Review stock at least twice a year
C	Dispose off the surplus items	Check and maintain the control	Review stock annually

XYZ-FSN classification exercise helps in the timely prevention of obsolescence.

XYZ Classification	**FSN Classification**		
	F	**S**	**N**
X	Light inventory control	Reduce stock to very low level	Quick disposal of items at optimum price
Y	Normal inventory control	Low level of stocks	Should be disposed as early as possible.
Z	Can reduce clerical work by increasing stocks.	Low level of stocks	Can afford to dispose at lower prices.

4.9. EXERCISE

Questions

1. Why is inventory an important consideration for managers?

2. State the purpose of inventory control.

3. Why would not a company all the time store large quantities of inventory to avoid shortages and stockouts?
4. Describe the major decisions that must be made in inventory control.

5. State some of the assumptions made in EOQ.

6. Discuss in detail the major inventory costs that are used in EOQ.

7. Classify the following 14 items in ABC categories.

Code Number	Monthly Consumption (Rs.)	Code Number	Monthly Consumption (Rs.)
D-179-0	451	D-196	214
D-195-0	1,052	D-198-0	188
D-186-0	205	D-199	172
D-191	893	D-200	170
D-192	843	D-204	5,056
D-193	727	D-205	159
D-195	412	D-212	3,424

How the policies with regard to safety stocks, order quantity, material control and inventory system will be different for the items classified as A, B and C?

8. The following information is provided for an item:
Annual demand: 12, 000 units; Ordering Cost: Rs. 60/order
Carrying Cost: 10%; Unit cost of item: Rs. 10 and Lead Time: 10 days.
There are 300 working days in a year. Determine EOQ and number of orders per year. In the past two years the demand rate has gone as high as 70 units/day. For a reordering system based on the inventory level, what should be the buffer stock? What should be the reorder level at this buffer stock? What would be the carrying costs for a year?

9. The demand for an item in a company is 18,000 units per year, and the company can produce the item at a rate of 3,000 per month. The cost of one set-up is Rs. 500 and the holding cost of one unit per month is 15 paise. The shortage cost of one unit is Rs. 240 per year. Determine the optimum manufacturing quantity and the number of shortages. Also determine the manufacturing time and the time between set-ups.

10. A product is sold at the rate of 50 pieces per day and is manufactured at a rate of 250 pieces per day. The set-up cost of the machines is Rs.1, 000 and the storage cost is found to be Re 0.0015 per piece per day. With labour charges of Rs. 3.20 per piece, material cost at Rs.2.10 per piece and overhead cost of Rs.4.10 per piece, find the minimum cost batch size if the interest charges are 8% (assume 300 working days in a year). Compute the optimal number of cycles required in a year for the manufacture of this product.

xxx-----xxx-----xxx

munotes.in

CHAPTER 5

SIMULATION TECHNIQUES

UNIT STRUCTURE

5.1 Introduction
5.2 Types of Simulation
5.3 Steps of Simulation Process
5.4 Advantages and Disadvantages of Simulation
5.5 Stochastic Simulation and Random Numbers
5.6 Monte Carlo Distribution
5.7 Simulation of Inventory Problems
5.8 Simulation of Queuing Problems
5.9 Applications of Simulation
5.10 Exercise

munotes.in

LEARNING OBJECTIVES

After studying this chapter, you should be able to

- Make distinction between analytical and simulation models.
- Appreciate what simulation is and what we can do with it.
- Know several definitions of simulation and appreciate the importance of simulation modelling.
- Understand advantages and disadvantages of simulation.
- Apply Monte Carlo simulation techniques for solving various types of problems.
- Develop random number intervals and use them to generate outcomes.
- Know several types of computer languages that are helpful in the simulation process.
- Know certain causes of simulation analysis failure and how to avoid them.

5.1. INTRODUCTION

Mathematical models cannot be manipulated to destroy its acceptability as a reasonable representation of a system under study. However, a model once constructed may be used to predict consequences of taking alternative actions. In particular, we could 'experiment' on the model by 'trying' alternative actions or parameters and compare their consequences. This 'experimentation' allows to answer 'what if' questions relating the effects of assumptions on the model response. Comparing the consequences of substituting various parameters into the model is referred to as simulating the model.

As the complexity of a model increases, simulation seeks to replicate the uncertainty in the model and assess model's response to events made to occur (i.e. simulated) with a frequency characterised by pre-specified probabilities.

Simulation is used in almost all fields, restricted only by our imagination and our ability to translate such imagination into a set of computer directives. The following are some simple examples which will make us appreciate what simulation is and what we can do with it.

Location of ambulances

Often, it is not known exactly where the demand for ambulance services will arise and, therefore, simulation is required to test alternative scheduling rules of the ambulances, their locations (in a geographical area), the response time to an emergency call and, of course, the overall service quality and the costs incurred if ambulances were to be configured in a certain way (in types, quantity, location, scheduling and staffing).

Design of computer systems

In general, it is difficult to specify future needs. Simulation is useful to highlight the basic characteristics of a computer system configuration (its speed, size, computing qualities, memory and so on.) in terms of the equipment used, the management of the computer system, the costs of such configurations and the resulting computational service which it provides to users.

Shop floor Management

The management of shop floor in factories is extremely complex because it involves many persons, machines, products parts, robots, material handling systems, Televisions and computer aided vision and inspection systems. Even if we were able to manage some specific aspects of the shop floor, integrating it into a simple mathematical model is very difficult, provided we can construct a model (however complex it may be) of the shop floor. Simulation is useful t analyse how a model could respond to certain policies (such as scheduling jobs to be performed on the basis of their due-dates, with earliest dates jobs performed first, by comparing alternative configuration and routing control policies, and so on). Simulation can also be used to test alternative work flows, the introduction of a new machinery, and in some cases the use of completely new manufacturing and management technologies (such as Just-In-Time manufacturing, flexible manufacturing, etc.) In such cases, simulation provides insights regarding the 'new' shop floor before it is implemented,

i.e. it provides, an experience and a learning tool which will give us greater confidence in the productivity and the management of the future system, at a relatively small cost.

Design of Optimal Replenishment Policy

In inventory control, the problem of optimal replenishment policy arises due to the probabilistic (stochastic) nature of demand and lead time. Thus, instead of manually trying out an appropriate replenishment alternative for each level of demand and lead time for a specific period and then selecting the best one, we process data (called experiment) on the computer and obtain the results in a very short time at a very small cost.

Design of a Queuing System

In queuing theory, the problem of balancing the cost of waiting against the cost of idle time of service facilities in the system arises due to the probabilistic nature of the inter-arrival times of customers and the time taken to complete the service to the customer. Thus, instead of trying out in actual manually with data to design a queuing system, we process the data on computers and obtain the expected value of various characteristics of the queuing system such as idle time of servers, average waiting time, queue length etc.

Other problems such as aircraft management routing, scheduling of bank tellers and location of bank branches, the deployment of fire stations, routing and dispatching when roads are not secured (where materials sent might not, potentially, reach their destination), the location and the utilisation of recreational facilities (such as parks, public swimming pools etc.) and many other problems have been studied or could be studied through simulation.

SIMULATION DEFINED

Simulation is one of the oldest operations research technique known to man - that is, the representation of the real world by numbers and other symbols that can be readily manipulated. To gain a better grasp of the real world, games such as chess to simulate the battles, backgammon to simulate racing and other games to simulate hunting and diplomacy were already invented in antiquity.

Today, a modern game like monopoly simulates the competitive arena of real estate. Many have played baseball with a deck of cards which has hits, strikeouts and walks with a cardboard, diamond and plastic chips as runners. The distribution of hits, runs and outs etc., in a deck of cards serves as a realistic reflection of the overall average with which each will occur in real life. Other games people play, generate experiences, understanding and gain knowledge which is basically what we will be trying to do through simulation.

What is new in modern simulation? It is the availability of computers which make it possible for us to deal with an extraordinary large quantity of details which can be incorporated into a model and the ability to manipulate the model over many 'experiments' (i.e. replicating all the possibilities that may be imbedded in the external world, and events would seem to recur). The modern use of the word simulation can be traced to the mathematicians Von Neumann and Ulam in the late 1940s when they developed the term *Monte Carlo analysis* while trying

first to 'break' the *Casino* at *Monte Carlo* (which they did not succeed!) and subsequently, applying it to the solution of nuclear shielding problems that were either too expensive for physical experimentation, or too complicated for treatment by known mathematical techniques.

An agreed definition for the world simulation has not been reached so far, however, a few definitions are stated as:

a. A simulation of a system or an organism operation of a model or simulator which is a representation of the system or organism. The model amenable to manipulation which would be impossible, too expensive or unpractical to perform on the entity it portrays. The operation of the model can be studied and for it, properties concerning the behaviour of the actual system can be inferred.

This definition is broad enough to be applied equally to military war games, business games, economic models etc. In this view simulation involves logical and mathematical constructs that can be manipulated on a digital computer using iterations or successive trials.

- Simulation is the process of designing a model of a real system and conducting experiments with this model for the purpose of understanding the behaviour (with the limits imposed by a criterion or set of criteria) for the operation of the system.
- Simulation is a numerical technique for conducting experiments on a digital computer, which involves certain types of mathematical and logical relationships necessary to describe the behaviour and structure of a complex real-world system over extended periods of time.

- Naylor et al.

Few other definitions of simulations are as under:

- 'X simulated Y' is true if and only if
  i. X and Y are formal systems.
  ii. Y is taken to be the real system.
  iii. X is taken to be an approximation to the real system, and
  iv. The rules of validity in X are non-error-free, otherwise X will become the real system.
- Simulation is the use of a system model that has the designed characteristics of reality in order to produce the essence of actual operation.

For Operations Research, simulation is a problem solving technique which uses a computer-aided experimental approach to study problems that cannot be analysed using direct and formal analytical methods. As a result, simulation can be thought of as a last resort technique. It is not a technique which should be applied in all cases. However, table 1 highlights what simulation is and what it is not.

Simulation - what it is/not

It is	It is not
- A technique which uses computers. - An approach for reproducing the processes by which events of chance and change are created in a computer.	- An analytical technique which provides exact solution. - A programming language but it could be programmed into a set of commands which can form a language to facilitate the

- A procedure for testing and experimenting on models to answer what if....then so and so....types of questions.	programming of simulation.

5.2. TYPES OF SIMULATION

There are several types of simulation. Few of them are listed below:

1. Deterministic/probabilistic simulation

The deterministic simulation is used when a process is very complex or consists of multiple stages with complicated (but known) procedural interactions between them. The measures of performance of such a system would be extremely detailed and time consuming. Formulating such a process as a simulation with fixed procedures (algorithms) allows the determination of the outcome and measure of performance in a straight forward manner.

In probabilistic simulation, one or more of the independent variables (e.g., the arrival rate of customers at a service-window is probabilistic, that is, it follows a certain probability distribution.

2. Time dependent and time independent simulation

In time independent simulation it is not important to know exactly when the event is likely to occur. For example, in inventory control situation, we may know that the demand is three units per day, but it is not necessary to know when during the day the item was demanded. On the other hand, in time dependent simulation it is important to the precise time when the event is likely to occur. For example, in queuing situation the precise time of arrival should be known (to know if the customer will have to wait).

3. Visual interactive simulation

Visual interactive simulation uses computer graphics displays to present the consequences of change in the value of input variation in the model. The decisions are implemented interactively while the simulation is running. These simulations can show dynamic systems that evolve over time in terms of animation. The decision maker watches the progress of the simulation in an animated form on a graphics terminal and can alter the simulation as it progresses.

4. Business Games

Business game simulation model involves several participants who need to play a role in a game that simulates a realistic competitive situation. Individuals or teams compete to achieve their goal, such as profit maximisation, in competition or cooperation with the other individuals or teams. The few advantages of business games are:

i. Participants learns much faster and the knowledge and experience gained are more memorable than passive instruction.

ii. Complexities, interfunctional dependencies, unexpected events, and other such factors can be introduced into the game to evoke special circumstances.

iii. The time compression - allowing many years of experience in only minutes or hours 0 lets that participants try out actions that they would not be willing to risk in an actual situation and see the result in the future.
iv. Provide insight into the behaviour of organisation. They dynamics of team decision-making style highlight the roles assumed by individuals on the teams, the effect of personality types and managerial styles, the emergence of team conflict and cooperation, and so on.

5. Corporate and financial simulations

The corporate financial simulation is used in corporate planning, especially the financial aspects. The models integrate production, finance, marketing, and possibly other function into one model either deterministic or probabilistic when risk analysis is desired.

5.3. STEPS OF SIMULATION PROCESS

The process of simulating a system consists of following steps:

1. Identify the problem

The simulation process is used to solve any problem only when the assumptions required for analytical models are not satisfied or there is no appropriate model developed for a system under study. For example, a queuing situation may be of interest but the arrival and/or service pattern do not meet the assumptions required to use queuing theory.

2. Identify the decision variables and decide the performance criterion (objective)

In the context of an inventory control situation, the demand (consumption rate), lead time and safety stock are identified as decision variables. These variables shall be responsible to measure the performance of the system in terms of total inventory cost under the decision rule - when to order.

3. Construct a simulation model

For developing a simulation model, an intimate understanding of the relationships among the elements of the system being studied is required. For this purpose, the influence diagram (drawn in a variety of different ways) is useful because simulation models for each of these diagrams may be formulated until one seems better or more appropriate than the other. Even after one has been chosen, it may be modified again and again before a final version is acceptable.

4. Testing and validating the model

Any simulation model must represent the system under study. This requires comparison of the model with the actual system - a validation process. A validated model should behave similar to the system under study. However, this validation condition may not be sufficient on model's predictive abilities.

The validation process requires

1. Determine whether the model is internally correct in a logical and programming sense called internal validity and
2. Determine whether it represents the system under study called external validity.

The first step involves checking the equations and procedures in the model for accuracy, both in terms of mistakes (or errors) and in terms of properly representing the system under study. The verification of internal validity can be simplified if the model is developed in modules and each module is tested as it is developed. Focusing on a particular module rather than trying to evaluate the logic of the entire model all at once facilitates identifying the source of errors and correcting them.

After verifying internal validity, the model is tested by substituting historical values into the model and seeing if it replicates what happens in reality. If the model passes this test, extreme values of the input variables are entered and the model is checked for the expected output.

A visual display of simulation results gives a better feel for what is happening in the model. The decision maker can make changes in the assumptions or input data and see the effect on the outputs. This improves the validity testing process.

5. Designing of the experiment

Experimental design refers to controlling the conditions of the study, such as the variables to include. This is in contrast to situations where observations are taken but the conditions of the study are not controlled.

It requires to determine factors considered fixed and variable in the model, levels of the factors to use, what the resulting dependent measures are going to be, how many times the model will be replicated and length of time of each replication and so on. For example, in a queuing simulation we may consider arrival and service rates constant but vary the number of servers and the evaluate the customer waiting times (dependent variable).

6. Run the simulation Model

Run the model on the computer to get the results in the form of operating characteristics.

7. Evaluate the Results

Examine the results of problem as well as their reliability and correctness. If the simulation process is complete, then select the best course of action (or alternative) otherwise make desired changes in model decision variables, parameters or design and return to step 3.

5.4. ADVANTAGES AND DISADVANTAGES OF SIMULATION

Advantages

The increased acceptance of simulation at various managerial levels is due to a number of following factors:

1. This approach is suitable to analyse large and complex real-life problems which cannot be solved by usual quantitative methods.

2. It is useful for sensitivity analysis of complex systems. In other words, it allows the decision-maker to study the interactive system variables and the effect of changes in these variables on the system performance in order to determine the desired one.

3. Simulation experiments are done with the model, not on the system itself. It also allows to include additional information during analysis that most quantitative models do not permit. In other words, simulation can be used to 'experiment' on a model of a real situation without incurring the costs of operating on the system.

4. Simulation can be used as a pre-service test to try out new policies and decision rules for operating a system before running the risk of experimentation in the real system.
5. The only 'remaining tool' when all other techniques become intractable or fail.

Disadvantages

The primary reasons of not often using simulation are as under:
1. Sometimes simulation models are expensive and take a long time to develop. For example, a corporate planning model may take a long time to develop and prove expensive also.
2. It is the trial and error approach that produces different solutions in repeated runs. This means it does not generate optimal solutions to problems.
3. Each application of simulation is ad hoc to a great extent.
4. The simulation model does not produce answers by itself. The user has to provide all the constraints for the solutions which he wants to examine.

5.5. STOCHASTIC SIMULATION AND RANDOM NUMBERS

In simulation, probability distributions are used to quantify outcomes in numerical terms by assigning a probability to each of possible outcomes. For example, if you flip a coin, the set of possible outcomes is {H, T}. A random variable assigns a number to the possible occurrence of each outcome. In simulation random variables are controlled numerically and are used to simulate elements of uncertainty which are defined in a model. This is done by generating (using the computer) outcomes with the same frequency as those encountered in the process being simulated. In this manner many experiments (also called simulation runs) can be performed, leading to a collection of outcomes that have a frequency (probability) distribution similar to that of the model under study.

To use simulation, it is necessary to learn generating the sample random events that make up the model. This helps to use the computer to reproduce the process through which chance is generated in actual situation. Thus a problem which involves many inter relationships among random variables can be evaluated as a function of given parameters. Thus process generation (simulating chance processes) and modelling are the two fundamental techniques of that are needed in simulation.

The first elementary and important type of process is the random process, which requires the selection of samples (or events) from a given distribution so that repetition of this selection process will yield a frequency distribution of sample values that matches the original distribution. When these samples are generated through some mechanical or electronic device - called pseudo random numbers. Alternately it is possible to use table of random numbers where the selection of number in any consistent manner will yield numbers that behave as if they were drawn from a uniform distribution.

Random numbers can also be generated using random number generator (which are inbuilt feature of spread sheets and many computer languages) tables, a roulette wheel etc.

Random numbers between 00 and 99 are used to obtain values of random variables that have a known discrete probability distribution in which the random variable of interest can assume one of a finite number of different values. In some applications, however, the random variables are continuous, that is, they can assume any real value according to a continuous probability distribution. For example, in queuing theory applications, the amount of time a server spends with a customer might follow an exponential distribution.

5.6. MONTE CARLO DISTRIBUTION

The Monte Carlo simulation technique involved conducting repetitive experiments on the model of the system under study with some known probability distribution to draw random samples (observations) using random numbers. If a system cannot be described by a standard probability distribution such as normal, Poisson, exponential, gamma etc. an empirical probability distribution can be constructed. The Monte Carlo simulation technique consists of following steps:

1. Setting up a probability distribution for variables to be analysed.
2. Building a cumulative probability distribution for each random variable.
3. Generate random numbers and then assign an appropriate set of random numbers to represent value or range (interval) of values for each random variable.
4. Conduct the simulation experiment using random sampling.
5. Repeat Step 4 until the required number of simulation runs has been generated.
6. Design and implement a course of action and maintain control.

Random Number Generation

Monte Carlo Simulation requires the generation of a sequence of random numbers. This sequence of random numbers helps in choosing random observations (samples) from the probability distribution.

Arithmetic Computation

The nth random number r_n consisting of k-digits generated by using multiplicative congruential method is given by

$$r_n \equiv p.r_{n-1} \text{ (modulo m)}$$

where p and m are positive integers, $p < m$, r_{n-1} is a k-digit number and modulo m means that r_n is the remainder when $p.r_{n-1}$ is divided by m. This means r_n and $p.r_{n-1}$ differ by an integer multiple of m. To start the process of generating random numbers, the first random number (also called seed) r_0 is specified by the user. Then using above recurrence relation a sequence of k-digit random number with period $h < m$ at which point the number r_0 occurs again can be generated.

For illustration, let p = 35, m = 100 and arbitrarily start with $r_0 = 57$. Since m - 1 = 99 and is the 2-digit number, therefore, it will generate 2-digit random numbers:

$r_1 = p.\ r_0$ (modulo m) = 35 x 57 (modulo 100) = 1, 995/100 = 95, remainder

$r_2 = p.\ r_1$ (modulo m) = 35 x 95 (modulo 100) = 3, 325/100 = 25, remainder

$r_3 = p.\ r_2$ (modulo m) = 35 x 25 (modulo 100) = 875/100 = 75, remainder

The choice of r_0 and p for any given value of m require great care, and the method used is also not a random process because sequence of numbers generated is determined by the input data for the method.

Thus, the numbers generated through this process are pseudo random numbers because there are reproducible and hence, not random.

The above defined recurrence relation can also be used to generate random numbers as decimal fraction between 0 and 1 with a desired number of digits. For this, the recurrence relations $u_n = r_n/m$ is used to generate uniformly distributed decimal fraction between 0 and 1.

Computer generator

The random numbers that are generated by using computer software are uniformly distributed decimal fractions between 0 and 1. The software works on the concept of cumulative distribution function for the random variables for which we are seeking to generate random numbers.

For example, for the negative exponential function with density function $f(x) = \lambda e^{-}\lambda x, 0< x < \infty$, the cumulative distribution function is given by

$$F(x) = \int_0^x \lambda e - \lambda x \, dx = 1 - e^{-}\lambda x$$

$$e^{-}\lambda x = 1 - F(x) \text{ or}$$

Taking logarithm on both sides, we have

$$-\lambda x = \log [1- F(x)]$$

$$x = -(1/\lambda) \log [1- F(x)] \text{ or}$$

If r = F(x) is a uniformly distributed random decimal fraction between 0 and 1, then the exponential variable associated with r is given by

$$x_n = -(1/\lambda) \log (1- r) = -(1/\lambda) \log r.$$

This is an exponential process generator since 1-r is a random number and can be replaced by r.

5.7. SIMULATION OF INVENTORY PROBLEMS

Examples

1. Using random numbers to simulate a sample, find the probability that a packet of 6 products does not contain any defective product, when the production line produces 10 per cent defective products. Compare your answer with the expected probability.

Solution

Given that 10 percent of the total production is defective and 90 per cent is non-defective. If we have 100 random numbers (0 to 99), then 90 or 90 percent of them represent non-defective products and remaining 10 or 10 percent of them represent defective products. Thus, the random numbers 00 to 89 are assigned to variables representing non-defective products and 90 to 100 are assigned to variables representing defective products.

If we choose a set of 2-digit random numbers in the range 00 to 99 to represent a packet of 6 products as shown below, then we would expect that 90 percent of the time they would fall in the range 00 to 89.

Sample Number	Random Number					
A	86	02	22	57	51	68
B	39	77	32	77	09	79
C	28	06	24	25	93	22
D	97	66	63	99	61	80
E	69	30	16	09	05	53
F	33	63	99	19	87	26
G	87	14	77	43	96	43
H	99	53	93	61	28	52
I	93	86	52	77	65	15
J	18	46	23	34	25	85

Here it may be noted that out of ten simulated samples 6 contain one or more defectives and 4 contain no defectives. Thus, the expected percentage of non-defective products is 40 percent. However, theoretically the probability that a packet of 6 products containing no defective product is $(0.9)^6 = 0.53144 = 53.14\%$.

2. A bakery keeps stock of a popular brand of cake. Previous experience shows the daily demand pattern for the item with associated probabilities, as given below:

Daily demand (number)	:	0	10	20	30	40	50
Probability	:	0.01	0.20	0.15	0.50	0.12	0.02

Use the following sequence of random numbers to simulate the demand for next 10 days.

Random Numbers: 25, 19, 65, 76, 12, 05, 73, 89, 19, 49.

Also estimate the daily average demand for the cakes on the basis of simulated data.

Solution:

Using the daily demand distribution, we obtain a probability distribution as shown in table 1.

Daily Demand Distribution

Daily Demand	Probability	Cumulative Probability	Random Number Interval
0	0.01	0.01	00
10	0.20	0.21	01 - 20
20	0.15	0.36	21 - 35
30	0.50	0.86	36 - 85
40	0.12	0.98	86 - 97
50	0.02	1.00	98 - 99

Conduct the simulation experiment for demand by taking a sample of 10 random numbers from a table of random numbers, which represent the sequence of 10 samples. Each random sample number here is a sample of demand.

The simulation calculations for a period of 10 days are given in the following table.

Simulation Experiments

Days	Random Number	Demand
1	40	30
2	19	10
3	87	40
4	83	30
5	73	30
6	84	30
7	29	20
8	09	10
9	02	10
10	20	10

	Total	220
	Expected Demand	220/10 = 22 units per day

5.8. SIMULATION OF QUEUING PROBLEMS

Examples

1. S dentist schedules all his patients for 30-minute appointments. Some of the patients take more or less than 30 minutes depending on the type of dental work to be done. The following summary shows the various categories of work, their probabilities and time actually needed to complete the work:

Category of Service	Time Required (Minutes)	Probability of Category
Filling	45	0.40
Crown	60	0.15
Cleaning	15	0.15
Extraction	45	0.10
Check-up	15	0.20

Simulate the dentist's clinic for 4 hours and determine the average waiting time for the patients as well as the idleness of the doctor. Assume that all the patients show up at the clinic at exactly their scheduled arrival time starting at 8.00 a.m. Use the following random numbers for handling the above problem:

40 82 11 34 25 66 17 79

Solution:

The cumulative probability distribution and random number interval for service time are shown in table 1.

Category of Service	Service Time Required (minutes)	Probability	Cumulative Probability	Random Number Interval
Filling	45	0.40	0.40	00 - 39
Crown	60	0.15	0.55	40 - 54
Cleaning	15	0.15	0.70	55 - 69

Extraction	45	0.10	0.80	70 - 79
Check-up	15	0.20	1.00	80 - 99

The various parameters of a queuing system such as arrival pattern of customers, service time, waiting time in the context of the given problem are shown in the table:

Arrival Pattern and Nature of Service

Patient Number	**Scheduled Arrival**	**Random Number**	**Category of Service**	**Service Time (Minutes)**
1	8.00	40	Crown	60
2	8.30	82	Check-up	15
3	9.00	11	Filling	45
4	9.30	34	Filling	45
5	10.00	25	Filling	45
6	10.30	66	Cleaning	15
7	11.00	17	Filling	45
8	11.30	79	Extraction	45

Computation of Arrivals, Departures and Waiting of Patients

Time	Event (Patient Number)	Patient Number (Time to Exit)	Waiting (Patient Number)
8.00	1 arrive	1 (60)	---
8.30	2 arrive	1(30)	2
9.00	1 departs; 3 arrive	2(15)	3
9.15	2 depart	3(45)	---
9.30	4 arrive	3(30)	4
10.00	3 departs; 5 arrive	4(45)	5
10.30	6 arrive	4(15)	5, 6
10.45	4 depart	5(45)	6
11.00	7 arrive	5(30)	6, 7
11.30	5 depart; 8 arrive	6(15)	7, 8
11.45	6 depart	7(45)	8
12.00	End	7(30)	8

The dentist was not idle during the entire simulated period. The waiting times for the patients were as follows:

Computation of Average Waiting Time

Patient	Arrival Time	Service Starts at	Waiting time (minutes)
1	8.00	8.00	0
2	8.30	9.00	30
3	9.00	9.15	15
4	9.30	10.00	30
5	10.00	10.45	45
6	10.30	11.30	60
7	11.00	11.45	45
8	11.30	12.30	60
			280

The average waiting time = 280/8 = 35 minutes.

5.9. APPLICATIONS OF SIMULATION

There is a wide range of applications of computer-based simulation models because it is an approach rather than an application of specific techniques.

The major use of computer-based Monte Carlo simulation model has been in the solution of complex queuing problems.

A number of job shop simulation programmes have been developed involving deterministic times for the individual operations of a given order. Due to different processing times for similar operations and different order operations sequences, it is difficult to predict the waiting time for a particular job at any given work centre. For better scheduling, orders must be scheduled with a provision of waiting at the various work centres they will pass through. Simulation can help in estimating accurately such waiting times.

A good deal of work has been done in the development of inventory simulation models such as determination of optimal reorder level and lot size under conditions of probabilistic demand and lead time, optimal review period and ordering policy for continuous review inventory models.

A number of network simulation models have also been developed. For example, with a randomly selected activity times the critical path can be evaluated. Repeating this process many times, the probability distribution of project completion time can be obtained as well as the probability that each given activity is on the critical path.

A variety of other problems which could be solved by simulation include:
1. Military studies of logistics, support planning and weapon systems effectiveness.
2. Studies of individual and group behaviour.
3. Financial studies involving risky investments.
4. Testing of decision rules for hospital admission and operating policies.

5.10. EXERCISE

Questions

1. Distinguish between solutions derived from simulation models and solution derived from analytical models?

2. What are random numbers? Why are random numbers useful in simulation models and solutions derived from analytical models?

3. What is Monte Carlo Simulation? Describe the idea of experimentation (Random Sampling) in simulation.

4. Describe the kind of problems for which Monte Carlo will be an appropriate method of solution.

5. Discuss the applications of simulation.

6. A book store wishes to carry a particular book in stock. Demand is not certain and there is a lead time of 2 days for stock replenishment. The probabilities of demand are given below:

Demand (units/day)	:	0	1	2	3	4
Probability	:	0.05	0.10	0.30	0.45	0.10

Each time an order is placed, the store incurs an ordering cost of Rs.10 per order. The store also incurs a carrying cost of Rs.0.5 per book per day. The inventory carrying cost is calculated on the basis of stock at the end of each day. The manager of the book store wishes to compare two options for his inventory decision.

A: Order 5 books when present inventory plus any outstanding order falls below 8 books.

B: Order 8 books when present inventory plus any outstanding order falls below 8 books.

Currently (beginning of 1st day) the store has a stock of 8 books plus 6 books ordered two days ago and are expected to arrive next day. Carryout simulation run for 10 days to recommend an appropriate option. You may use random numbers in the sequences. Using first number for day one.

89, 34, 78, 63, 61, 81, 19, 16, 13, 73.

7. A company trading in motor vehicle spare parts wishes to determine the levels of stock it should carry for the items in its range. Demand is not certain and there is a lead time for stock replenishment. For an item A, the following information is obtained:

Demand (units/day)	:	3	4	5	6	7
Probability	:	0.10	0.20	0.30	0.30	0.10
Carrying cost (per unit/day)	:	Rs. 2				
Ordering Cost (per order)	:	Rs. 50				
Lead Time for replenishment	:	3 days				

Stock on hand at the beginning of the simulation exercise was 20 units.

Carry out a simulation run over a period of 10 days with the objective of evaluating the inventory rule: Order 15 units when present inventory plus any outstanding order falls below 15 units.

You may use random numbers in the sequence of: 0, 9, 1, 1, 5, 1, 8, 6, 3, 5, 7, 1, 2, 9 using the first number for day one. Your calculation should include the total cost of operating this inventory rule for 10 days.

8. A firm has a single channel service station with the following arrival and service time probability distributions:

Inter arrival Time (minutes)	Probability	Service Time (minutes)	Probability
10	0.10	5	0.08
15	0.25	10	0.14
20	0.30	15	0.18

25 30	0.25 0.10	20 25 30	0.24 0.22 0.14

The customer's arrival at the service station is a random phenomenon and the time between the arrival varies from 10 minutes to 30 minutes. The service time varies from 5 minutes to 30 minutes. The queuing process begins at 10.00 a.m. and proceeds for nearly 8 hours. An arrival goes to the service facility immediately, if it is free. Otherwise it will wait in a queue. The queue discipline is first-come-first-served.

If the attendant's wages are Rs.10 per hour and the customer's waiting time costs Rs. 15 per hour, then would it be an economical proposition to engage a second attendant? Answer using Monte Carlo simulation technique.

xxx-----xxx-----xxx

munotes.in

CHAPTER 6

QUEUING THEORY

UNIT STRUCTURE

6.1 Introduction
6.2 The structure of a Queuing System
6.3 Elements of Queuing System
6.4 Applications and Limitations
6.5 Operating Characteristics of Queuing System
6.6 Basic Characteristics of Queuing System
6.7 Service Channels
6.8 Multi- Server Queuing Models
6.9 Exercise

LEARNING OBJECTIVES

After studying this chapter, you should be able to

- identify and examine situations that generate queuing problems.
- describe the trade-off between cost of service and cost of waiting time.
- understand various components (or parts) of a queuing system and description of each of them.
- analyse a variety of performance measures (operating characteristics) of a queuing systems
- derive relationships among variety of performance measure using probabilistic distributions.
- understand distinction among several queuing models and derive performance measures for each of them.

munotes.in

6.1. INTRODUCTION

A common situation occurring in everyday life is that of queuing or waiting in a line. Queues are usually seen at bus stops, ticket booths, doctors' clinics, bank counters, traffic lights and so on. Queues are also found in workshops where the machines wait to be repaired; at a tool crib where the mechanics wait to receive tools, in a warehouse where items wait to be used, incoming calls wait to mature in the telephone exchange, trucks wait to be unloaded, airplanes wait either to take off or land and so on.

In general, a queue is formed at a queuing system when either customers (human beings or physical entities) requiring service wait due to number of customers exceeds the number of service facilities, or service facilities do not work efficiently and take more time than prescribed to serve a customer.

Queuing theory can be applied to a variety of operational situations where it is not possible to predict accurately the arrival rate (or time) of customers and service rate (or time) of service facility or facilities. In particular, it can be used to determine the level of service (either the service rate or the number of service facilities) that balances the following two conflicting costs:

i. Cost of offering the service

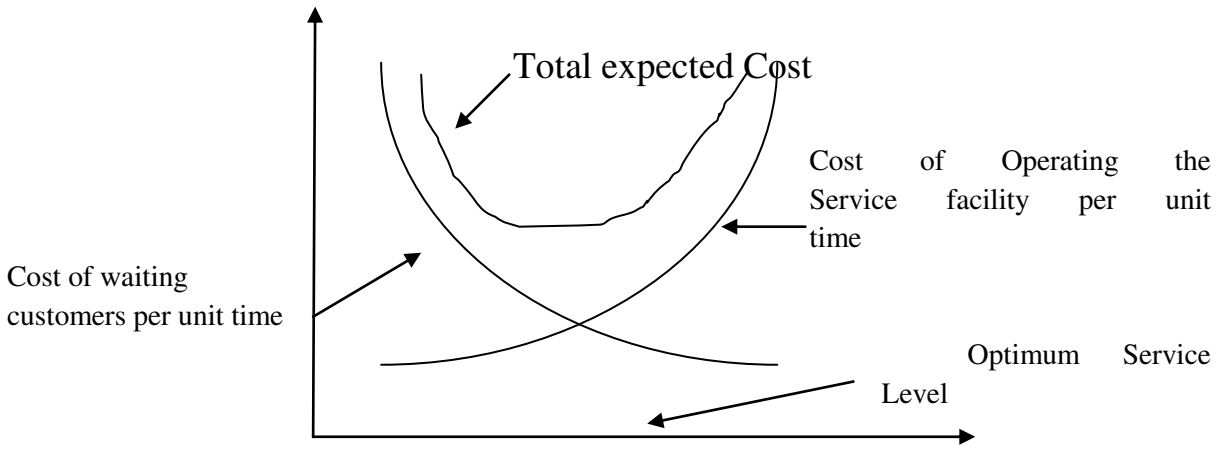
ii. Cost incurred due to delay in offering service

The first cost is associated with the service facilities and their operations and the second represents the cost of customer's waiting time.

Obviously, an increase in the existing service facilities would reduce the customer's waiting time. Conversely, decreasing the level of service should result in long queue(s). This means an increase (decrease) in the level of service increases (decreases) the cost of operating service facilities but decreases (increases) the cost of waiting. Figure 1 illustrates both types of costs as a function of level of service. The optimum service level is one that minimises the sum of the two costs.

Since cost of waiting is difficult to estimate, it is usually measured in terms of loss of sales or goodwill when the customer is a human being who has no sympathy with the service. But, if the customer is a machine waiting for repair, then cost of waiting is measured in terms of the cost of lost production.

Figure 1 Queuing Costs Vs Level of Service



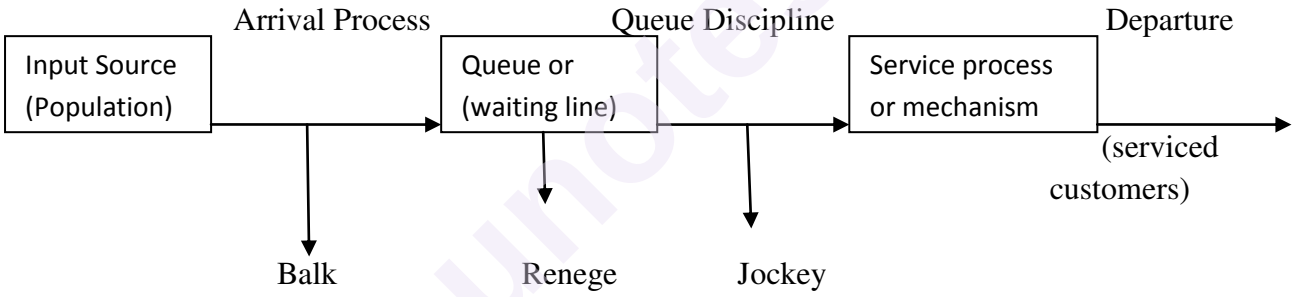
Many practical situations in which study of queuing theory can provide solution to waiting line problems are listed in the following table.

Situations	Customers	Service Facilities
Petrol Pumps	Automobiles	Pumps/Passionel
Hospital	Patients	Doctors/Nurses/Rooms
Airport	Aircraft	Runways
Post Office	Letters	Sorting System
Job Interviews	Applicants	Interviewers

6.2. THE STRUCTURE OF A QUEUING SYSTEM

A queuing system is composed of the following components (or parts) each of which is described below:

1. Calling population (or input source)
2. Queuing process
3. Queue discipline
4. Service Process (or mechanism)



Customers are defined as those in need of services. Customers requiring services are generated at different times by a calling population, commonly known as input source.

The manner in which customers arrive at the service facility individually or in batches, at scheduled or unscheduled time is called the arrival process. The customer's entry into the queuing system depends upon the queue conditions.

The service is provided by one or more service facilities. This may be person (bank teller, barber etc.). or a space (airport runway, hospital bed, parking lot etc.). Customers from a queue are selected for services according to certain rules known as queue discipline. A service facility may be without server (self-service) or include one or more servers operating either in a series (as a team or in parallel) (multiple service channels). The rate (constant or random) at which service is rendered is known as the service process. After the service is rendered, the customer leaves the system.

If server is idle at the time of customer's arrival, then customer is served immediately, otherwise the customer is asked to join a queue or waiting line, which may have single, multiple or ever priority lines. The order of service may be random, by appointment, first-cum-first served or some other order. The queue can also be considered to be either finite or infinite.

6.3. ELEMENTS OF QUEUING SYSTEMS

Population of Customers can be considered either limited (closed systems) or unlimited (open systems). Unlimited population represents a theoretical model of systems with a large number of possible customers (a bank on a busy street, a motorway petrol station). Example of a limited population may be a number of processes to be run (served) by a computer or a certain number of machines to be repaired by a service man. It is necessary to take the term "customer" very generally. Customers may be people, machines of various nature, computer processes, telephone calls, etc.

Arrival defines the way customers enter the system. Mostly the arrivals are random with random intervals between two adjacent arrivals. Typically the arrival is described by a random distribution of intervals also called Arrival Pattern.

Queue represents a certain number of customers waiting for service (of course the queue may be empty). Typically the customer being served is considered not to be in the queue. Sometimes the customers form a queue literally (people waiting in a line for a bank teller). Sometimes the queue is an abstraction (planes waiting for a runway to land). There are two important properties of a queue: Maximum Size and Queuing Discipline.

Maximum Queue Size (also called System capacity) is the maximum number of customers that may wait in the queue (plus the one(s) being served). Queue is always limited, but some theoretical models assume an unlimited queue length. If the queue length is limited, some customers are forced to renounce without being served.

Queuing Discipline represents the way the queue is organised (rules of inserting and removing customers to/from the queue). There are these ways:

1) FIFO (First In First Out) also called FCFS (First Come First Serve) - orderly queue.

2) LIFO (Last In First Out) also called LCFS (Last Come First Serve) - stack.

3) SIRO (Serve In Random Order).

4) Priority Queue, that may be viewed as a number of queues for various priorities.

5) Many other more complex queuing methods that typically change the customer's position in the queue according to the time spent already in the queue, expected service duration, and/or priority. These methods are typical for computer multi-access systems.

Most quantitative parameters (like average queue length, average time spent in the system) do not depend on the queuing discipline. That's why most models either do not take the queuing discipline into account at all or assume the normal FIFO queue. In fact the only parameter that depends on the queuing discipline is the variance (or standard deviation) of the waiting time. There is this important rule (that may be used for example to verify results of a simulation experiment):

The two extreme values of the waiting time variance are for the FIFO queue (minimum) and the LIFO queue (maximum).

Theoretical models (without priorities) assume only one queue. This is not considered as a limiting factor because practical systems with more queues (bank with several tellers with separate queues) may be viewed as a system with one queue, because the customers always select the shortest queue. Of course, it is assumed that the customers leave after being served. Systems with more queues (and more servers) where the customers may be served more times are called Queuing Networks.

Service represents some activity that takes time and that the customers are waiting for. Again take it very generally. It may be a real service carried on persons or machines, but it may be a CPU time slice, connection created for a telephone call, being shot down for an enemy plane, etc. Typically a service takes random time. Theoretical models are based on random distribution of service duration also called Service Pattern. Another important parameter is the number of servers. Systems with one server only are called Single Channel Systems, systems with more servers are called Multi Channel Systems.

Output represents the way customers leave the system. Output is mostly ignored by theoretical models, but sometimes the customers leaving the server enter the queue again ("round robin" time-sharing systems).

Queuing Theory is a collection of mathematical models of various queuing systems that take as inputs parameters of the above elements and that provide quantitative parameters describing the system performance.

Because of random nature of the processes involved the queuing theory is rather demanding and all models are based on very strong assumptions (not always satisfied in practice). Many systems (especially queuing networks) are not soluble at all, so the only technique that may be applied is simulation.

Nevertheless queuing systems are practically very important because of the typical trade-off between the various costs of providing service and the costs associated with waiting for the service (or leaving the system without being served). High quality fast service is expensive, but costs caused by customers waiting in the queue are minimum. On the other hand long queues may cost a lot because customers (machines e.g.) do not work while waiting in the queue or customers leave because of long queues. So a typical problem is to find an optimum system configuration (e.g. the optimum number of servers). The solution may be found by applying queuing theory or by simulation.

6.4. APPLICATIONS AND LIMITATIONS

Applications

The queueing theorie is used to analyze computer, telecommunication systems, traffic systems (traffic flue), logistic and manufacturing systems.

Some examples :

-Where customers are involved such as restaurants, supermarket, airport, …

-useful in manufactoring units

-applicable for the problem of machine breakdown and repair

-applicable for the sheduling of jobs in production control

-provide solution of inventory control problems

Limitations

-Most of the queueing models are still complex and are not easy to understand

-Many times form of theoratical distribution applicable to given queueing situations is not known

-The study of queueing problems become more difficult if the queueing discipline is not in "first in, first out".

6.5. OPERATING CHARACTERISTICS OF QUEUING SYSTEM

The operating characteristics of a queuing system refer to the numerical values of the probability distributions of various decision variables like arrival rate, number of facilities, service time, time length, priority system etc.

1. Queue Length

Probability distribution of queue length can be obtained with the help of the given probability distribution of the arrival and service process. A large queue indicates poor service facility or a need for more space. On the other hand, small queue indicates excess of service facilities.

2. Waiting time in queue

This is the time spent by a customer in the queue before the commencement of his service. Long waiting times may increase the customer's dissatisfaction and potential loss of future revenue.

3. Waiting time in system

Waiting time in a system is the total time spent by a customer in the queue plus the service time. Long waiting times may indicate a change in priority rules is needed.

6.6. BASIC CHARACTERISTICS OF A QUEUING SYSTEM

Even though there are many types of queuing systems, all such systems are fully defined and classified based on the following characteristics.

Input Process or Arrival Pattern of Customers

This is concerned with the pattern in which customers arrive and join the queuing system. The arrival of customers towards service station and the behaviour of customers in the queue are considered in the input process.

1. Arrival time distribution

The arrival of the customers to the service station is governed by certain probability loss. Generally, it is assumed that the arrival rate follows Poisson distribution with mean arrival rate, λ, or mean inter-arrival time has an exceptional distribution with mean $1/\lambda$. It is used to

count the number of arrivals per unit time. The number of customers may be from finite or infinite sources. Customers arrive at the service facility in batches of fixed size or of variable size or one by one. Sometimes, more than one arrival (in bulk or batches) can enter in the system simultaneously.

2. Behaviour of the customer

The reaction of a customer upon entering the system is called the behaviour of the customer. Customers generally behave in the following four ways:

a. Balking

When the customer comes to the service station he may find a lengthy queue already waiting to receive the service, so he may not join the queue. This is defined as balking.

b. Reneging

A customer who is waiting in the queue may leave the queue due to his impatience. This is defined as reneging.

c. Jockeying

A customer who is waiting in a queue may leave that queue and join with another queue, which is smaller in length. It is defined as jockeying.

d. Collusion

Two or more customers may go to the service station but only one customer will receive the service. This is collusion.

An arrival pattern, that does not change time and in which a steady state condition occurs, is called a stationary arrival time, whereas in an arrival pattern that is time dependent, the input is called a non-stationary arrival time.

6.7. SERVICE CHANNELS

Service facilities or service channels may be in series, parallel or mixed. Series arrangement includes a sequence of a number of service facilities, arranged such that a customer must go through one facility after another in a particular sequence. Each service facility works independently of others. Basically there are two structures of waiting line situations:

1. Single-channel service system
2. Multi-channel service system

1. Single-channel service system

There may be only one service station a server to provide service. This is defined as a single-channel service system. A single server model involves a single server and a single line of customers.

To further specify the model, we make the following assumptions:

1. The customer population is infinite and all customers are patient.
2. Customers arrive according to a Poisson distribution, with a mean arrival rate of λ.
3. The service distribution is exponential, with a mean service rate of μ.
4. Customers are served on a first-come-first basis.
5. The length of the waiting line is unlimited.

2. Multi-channel service system

Sometimes two or more service stations or servers may provide the same type of service, which is defined as a multi-channel service system. With a multiple server model, customers form a single line and choose one of *s* servers when one is available. The service system has only one phase. In addition to the assumptions for the single-server model or multi-server model, we assume that there are *s* identical servers, and the service distribution for each server is exponential, with a mean service time of $1/\mu$..

FOUR BASIC STRUCTURES OF WAITING LINE SITUATIONS

The time gap between the commencement of service and the completion of service time for a customer is called the service time for a customer.

a. Service time may be i. constant or ii. scattered.

b. Service time distribution may be i. Constant time, ii. Poisson, iii. Exponential and iv. Erlang.

Anyway, the most frequently used distribution to describe service time is the negative exponential distribution.

Capacity of the service station

Sometimes the service station may accommodate any number of customers (infinitely large). It is then called an infinite system. But sometimes, the service station may accommodate only a specific number of customers and the system is called a finite system.

Examples

1. A television repairman finds that the time spent on his jobs has an exponential distribution with a mean of 30 minutes. If he repairs sets in the order in which they came in, and if the arrival of sets follows a Poisson distribution approximately with an average rate of 10 per 8-hour day, what is the repairman's expected idle time each day? How many jobs are ahead of the average set just brought in?

<u>Solution:</u>

From the data of the problem, we have

$\lambda = 10/8 = 5/4$ sets per hour; and $\mu = (1/30)\ 60 = 2$ sets per hour

a. Expected idle time of repairman each day

Number of hours for which the repairman remains busy in an 8-hour day (traffic intensity) is given by

$(8)\ (\lambda/\mu) = (8)\ (5/8) = 5$ hours

Hence, the idle time for a repairman in an 8-hour day will be (8 - 5) = 3 hours.

b. Expected (or average) number of TV sets in the system

$L_s = \frac{\lambda}{\mu - \lambda} = \frac{5/4}{2-(\frac{5}{4})} = 5/3 = 2$(approx.) TV sets.

2. Arrivals at telephone booth are considered to be Poisson with an average time of 10 minutes between one arrival and the next. The length of phone call is assumed to be distributed exponentially, with mean 3 minutes.

a. What is the probability that a person arriving at the booth will have to wait?

b. The telephone department will install a second booth when convinced that an arrival would expect waiting for at least 3 minutes for a phone call. By how much should the flow of arrivals increase in order to justify a second booth?
c. What is the average length of the queue that forms from time to time?
d. What is the probability that it will take him more than 10 minutes altogether to wait for the phone and complete his call?

Solution:

From the data of the problem, we have

$\lambda = 1/10 = 0.10$ person per minute and $\mu = 1/3 = 0.33$ person per minute

a. Probability that a person has to wait at the booth.

$P(n > 0) = 1 - P_0 = \lambda/\mu = 0.10/0.33 = 0.3$

b. The installation of second booth will be justified only if the arrival rate is more than the waiting time.

Let λ' be the increased arrival rate. Then expected waiting time in the queue will be

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

$3 = \frac{\lambda'}{0.33(0.33 - \lambda')}$; or $\lambda' = 0.16$

where $W_q = 3$ (given) and $\lambda = \lambda'$ (say) for second booth.

Hence, the increase in the arrival rate is 0.16 - 0.10 = 0.06 arrivals per minute.

c. Average length of non-empty queue

$L = \frac{\mu}{\mu - \lambda} = 0.33/0.23 = 2$ customers (approx.)

d. Probability of waiting for 10 minutes or more is given by

$P(t \geq 10) = \int_{10}^{\infty} \frac{\lambda}{\mu}(\mu - \lambda)\, e^{-(\mu - \lambda)t} dt$

$= \int_{10}^{\infty} (0.03)(0.23))\, e^{-0.23\,t} dt = 0.069\, [e^{-0.23\,t}/- 0.23]_{10}\infty = 0.03$

This shows that 3 percent of the arrivals on an average will have to wait for 10 minutes or more before they can use the phone.

6.8. MULTI-SERVER QUEUING MODELS

Examples

1. A tax consulting firm has 4 service counters in its office to receive people who have problems and complaints about their income, wealth and sales taxes. Arrivals average 80 persons in an 8-hour service day. Each tax advisor spends an irregular amount of time servicing the arrivals which have been found to have an exponential distribution. The average service time is 20 minutes. Calculate the average number of customers in the system, average number of customers waiting to be serviced, average time a customer spends in the system, and average waiting time for a customer. Calculate how many hours each week does a tax adviser spend performing his job. What is the probability that a customer has to wait before he gets service? What is the expected number of idle tax advisors at any specified time?

Solution:

Given that $\lambda = 10$/hour; $\mu = 3$/hour, 4 and $\rho = \lambda/s\mu = 5/6$

a. Probability of no customer in the system,

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

$$= [\sum_{n=0}^{3} \frac{1}{n!} (10/3)^n + \frac{1}{4!}(10/3)^4 (12/12-10)]^{-1}$$

$$= [1 + 10/3 + 1/2 (10/3)^2 + 1/6(10/3)^4 12/12 - 10]^{-1}$$

$= 0.021$

b. Average number of customers in the system

$$L_s = L_q + \lambda/\mu = [\frac{1}{(s-1)!} (\lambda/\mu)^s \frac{\lambda\mu}{(s\mu-\lambda)2}] P_0 + \lambda/\mu$$

$$= [\frac{1}{3!} (10/3)^4 \frac{30}{(12-10)2}] \times 0.021 + 10/3 = 6.57$$

c. Average number of customers waiting in the queue (queue length),

$L_q = L_s - \frac{\lambda}{\mu} = 6.57 - (10/3) = 3.24$ customers

d. Average time a customer spends in the system

$W_s = W_q + 1/\mu = L_q/\lambda + 1/\mu = 3.24/10 + 1/3 = 0.657$ hour or 39.42 minutes

e. Average time a customer waits for service in the queue

$W_q = L_q/\lambda = 3.24/10 = 0.324$ hour or 19.44 minutes.

f. time spent by a tax counsellor, i.e. utilisation factor,

$\rho = \lambda/s\mu = 5/6 = 0.833$ hour or 50 minutes

The expected time spent in servicing customers during an 8-hour day is 8 x 0.833 = 6.66 hours. Thus, on an average, a tax advisor is busy for 6.66 x (40/8) = 33.30 hours based on a 40 hours' week.

g. Probability that a customer has to wait

$$P_w (n \geq s) = \frac{1}{s!}(\lambda/\mu)^s (s\mu/s\mu-\lambda) . \rho_0$$

$$= \frac{1}{4!}(10/3)^4 (4 \times 3/4 \times 3-10) (0.021) = 0.622$$

h. The expected number of idle advisers at any specified time can be obtained by adding the probability of 3 idle, 2 idle and 1 idle adviser. That is

Expected number of idle advisers $= 4 P_0 + 3P_1 + 2P_2 + P_3$

$= 4 (0.021) + 3(0.070) + 2(0.118) +0.131 - 0.661$

This means less than one 0.661 adviser is idle on an average at any instance of time.

6.9. EXERCISE

Questions

1. Derive the difference equations for the queuing model {(M/M/1) : (∞/FCFS)}. How would you proceed to solve the model?

2. In a single server, Poisson arrival and exponential service time queuing system show that probability P_n of n customers in steady-state satisfies the following equations:

$\lambda P_0 = \mu P_1$; n = 0

$(\lambda + \mu) P_1 = \mu P_2$; n = 1

$(\lambda + \mu) P_n = \mu P_{n+1} + \lambda P_{n-1}$; $n \geq 2$

3. At what average rate must a clerk at a super market work in order to ensure a probability of 0.90 so that the customer will not wait longer than 12 minutes? It is assumed that there is only one counter at which customers arrive in a Poisson fashion at an average rate of 15 per hour. The length of service by the clerk has an exponential distribution.

4. Consider a self-service store with one cashier. Assume Poisson arrivals and exponential service times. Suppose that nine customers arrive on the average every 5 minutes and the cashier can serve 10 in 5 minutes. Find

a. Average number of customers queuing for service.

b. Probability of having more than 10 customers in the system and

c. Probability that a customer has to queue for more than 2 minutes.

If the service can be speed up to 12 in 5 minutes by using a different cast register, what will be the effect on the quantities (a0, (b) and (c).

5. For the single server, finite (or limited) queuing system, find the

a. Average number of customers in the system and

b. Average queue length.

xxx----xxx----xxx

munotes.in