AKLC

ASHOK KUMAR LEARNING CENTER

Learning is Life .. !!

Live it .. !!

SSM

[VERSION-2]

Tutorial classes for B.E (CSE/ISE)

All semesters

www.aklcindia.com | office@aklcindia.com

9742013378 | 9742024066

Vijayanagar | Jayanagar

UNIT 1 : INTRODUCTION

Syllabus

* When simulation is appropriate tool and when it is not appropriate
* Advantages and Disadvantages of simulation
* Areas of Application
* Systems and System Environment
* Components of a System
* Discrete and Continuous Systems
* Model of a System
* Types of models
* Discrete Event Simulation System
* Steps in a Simulation Study

* Simulation Examples
* Simulation of Queuing Systems
* Simulation of Inventory systems

— 8 Hours.

* Simulation is the imitation of the operation of a real-world process or system over time. Whether done by hand or computer, simulation involves the generation of artificial history of a system and the observation of that artificial history to draw inferences concerning the operating characteristics of the real system.

* The behaviour of a system as it evolves over time is studied by developing a simulation model.

* This model usually takes the form of a set of assumptions concerning the operation of the system. These assumptions are expressed in mathematical, logical, and symbolic relationships b/w the entities (objects of interest) of the system.

* Simulation modelling can be used

1. As an Analysis tool for predicting the effect of changes to existing systems.
2. As an Design tool to predict the performance of new systems.

WHEN SIMULATION IS THE APPROPRIATE TOOL

1. Simulation enables the study of, and experimentation with, the internal interactions of a complex system.
2. Informational, organizational, and environmental changes can be simulated, and the effect of these alterations on the model's behavior can be observed.
3. The knowledge gained in designing a simulation model may be of great value toward suggesting improvements in the system under investigation.
4. By changing simulation inputs and observing the resulting outputs, valuable insight may be obtained into which variables are most important and how variables interact.
5. Simulation can be used as a pedagogical device to reinforce analytic solution methodologies.
6. Simulation can be used to experiment with new designs or policies prior to implementation, so as to prepare for what may happen.
7. Simulation can be used to verify analytic solutions.
8. By simulating different capabilities for a m/c, requirements can be determined.

9. Simulation models designed for training allow learning without the cost and disruption of on-the-job learning.

10. Animation shows a system in simulated operation so that the plan can be visualized.

11. The modern system is so complex that the interactions can be treated only through simulation.

WHEN SIMULATION IS NOT APPROPRIATE

1. When the problem can be solved using common sense.
2. When the problem can be solved analytically.
3. When it is easier to perform direct experiments.
4. When the simulation costs exceed ~~the~~ savings.
5. When the resources or time are not available.
6. When system behavior is too complex or can't be defined.
7. When there isn't the ability to verify and validate the model.

ADVANTAGES AND DISADVANTAGES OF SIMULATION.

ADVANTAGES

1. New policies, operating procedures, decision rules, information flows, organizational procedures and so on can be explored without disrupting the ongoing operations of the real system.
2. New hardware designs, physical layouts, transportation systems, and so on, can be tested without committing resources for their acquisition.
3. Hypothesis about how and why certain phenomena occur can be tested for feasibility.
4. Insight can be obtained about the interaction of variables.
5. Insight can be obtained about the importance of variables to the performance of the system.
6. Bottleneck analysis can be performed indicating where work-in process, information, materials, and so on are being excessively delayed
7. Simulation study can help in understanding how the system operates rather than how the individuals thinks the system operates.
8. "what-if" questions can be answered. This is particularly useful in designing the new system.

DISADVANTAGES

1. Model building requires special training.
   * It is an art that is learnt over time and experience
   * Furthermore, if two models are constructed by two competent individuals, they may have similarities, but it is highly unlikely that they will be the same.

2. Simulation results may be difficult to interpret
   * Since most simulation outputs are essentially random variables, it may be hard to determine whether an observation is a result of system interrelationships or randomness.

3. Simulation modelling and analysis can be time consuming and expensive.
   * Skimping on resources for modelling and analysis may result in a simulation model or analysis that is not sufficient for the task.

4. Simulation is used in some cases when an analytic solution is possible, or even preferable.
   * This might be particularly true in the simulation of some waiting lines where closed-form queuing models are available.

AREAS OF APPLICATION

1. Manufacturing Applications
2. Semi conductor Manufacturing
3. Construction Engineering and project management
4. Military Applications
5. Logistics, supply chains, and distributed applications
6. Transportation modes and traffic.
7. Business process simulation
8. Health care.

SYSTEMS AND SYSTEM ENVIRONMENT

* System

A system is defined as a group of objects that are joined together in some regular interaction or interdependent toward the accomplishment of some purpose.

* System Environment

A system is often affected by changes occuring outside the system. Such changes are said to occur in system environment.

COMPONENTS OF THE SYSTEM :

* **Entity**
  An object of interest in the system

* **Attributes**
  A property of an entity.

* **Activity**
  A time period of specified length.

* **State**
  The collection of variables necessary to describe the system at any time, relative to the objectives of the study.

* **Event**
  An instantaneous occurrence that may change the state of a system.

* **Endogeneous.**
  Used to describe activities and events occurring within a system.

* **Exogeneous**
  Used to describe activities and events in an environment that affects the system.

Examples of systems and components.

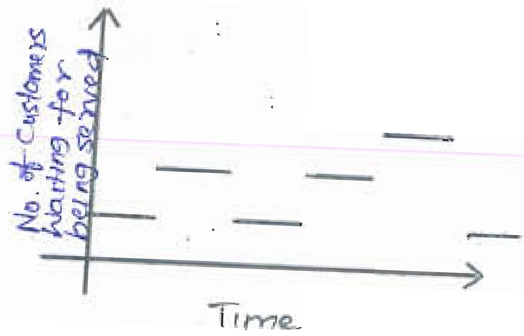
System	Entity	Attributes	Activities	Events	State Variables.
Banking	customers	checking account balance	Make deposits	Arrival, Departure	No. of Busy Tellers, No. of customers waiting
Production	Machines	speed, capacity	Welding, stamping	Breakdown	Status of machines [Busy, idle, down]

DISCRETE AND CONTINUOUS SYSTEMS

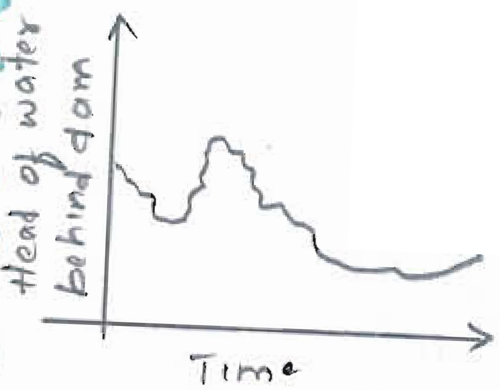
* **Discrete System:**

A discrete system is one in which the state variables change only at a discrete set of points in time

Ex: Bank.



* **Continuous system**



A continuous system is one in which the state variables change continuously over time

Ex: Head of water behind a Dam.

MODELS OF A SYSTEM

Model is a representation of a system for the purpose of studying the system.
It is the simplification of the system.

TYPES OF MODELS

1. Static or Dynamic Simulation Models
2. Deterministic or Stochastic Simulation Models.

Static v/s Dynamic simulation models

→ Static simulation model (called Monte carlo simulation) represents a system at a particular point in time.

→ Dynamic simulation model represents systems as they change over time.

Deterministic v/s Stochastic Simulation models

→ Deterministic simulation models contain no random variables and have a known set of inputs which will result in a unique set of outputs.

→ Stochastic simulation model has one or more random variables as inputs. Random inputs lead to random outputs.

DISCRETE EVENT SYSTEM SIMULATION

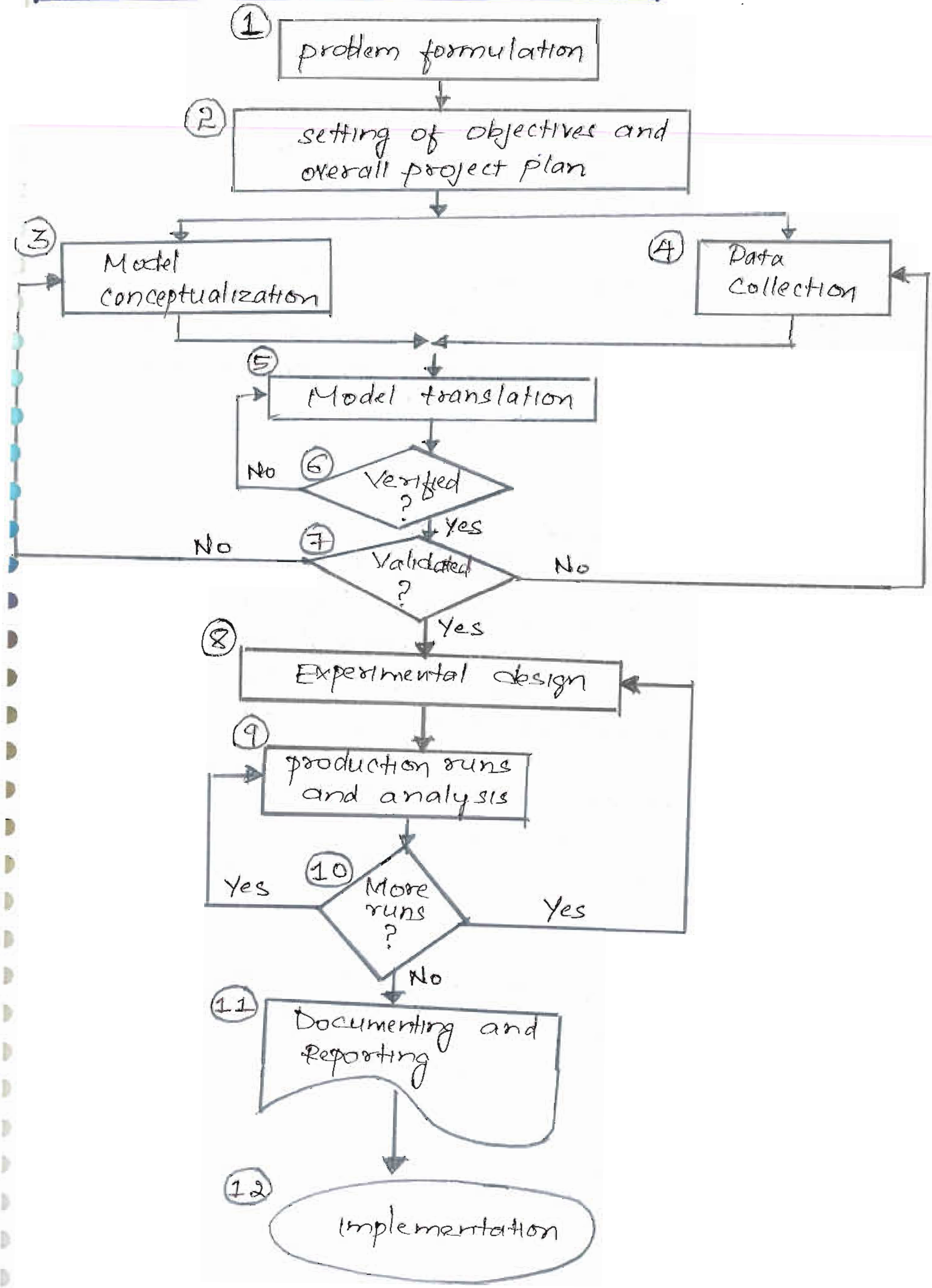
* Discrete Event system simulation is the modeling of systems in which the state variables changes only at a discrete set of points in time.

* The simulation models are analyzed by numerical rather than by analytical methods.

  → Analytical methods employ the deductive reasoning of mathematics to solve the model.

  → Numerical methods employ computational procedures to solve mathematical models.

STEPS IN A SIMULATION STUDY.



1. problem formulation.

* Every study begins with a statement of problem provided by policy ~~p~~ makers.
* Analyst ensures it is clearly understood. If it is developed by analyst, policy makers should understand and agree with it.

2. Setting of objectives and overall project plan.

* A determination must be made whether simulation is the appropriate methodology for problem formulated
* A statement of the alternatives systems to be considered.

3. Model Conceptualization

* Art of modeling is enhanced by an ability to abstract the essential features of an problem, to select and modify basic assumptions that characterize the system, and then to enrich and elaborate the model until a useful approximation results.
* Model conceptualization enhance the quality of the resulting model and increase the confidence of the model user in the application of the model.

4. Data Collection

* There is a constant interplay between the construction of the model and the collection of needed i/p data.
* Done in the early stages.
* objective kind of data are to be collected.
* As the complexity of the model changes, the required data elements may also change.

5. Model Translation.
* Real-world systems results in models that require a great deal of information storage and computation
* It can be programmed by using simulation languages or special purpose simulation softwares.
* Simulation languages are powerful and flexible.
* Simulation software models development time can be reduced.

6. Verified?
* Is the computer program performing properly?
* Debugging for correct input parameters and logical structures.

7. Validated?
* The determination that an model is an accurate representation of real system.
* Validation is achieved through calibration of the model.

8. Experimental Design.
* The design on the length of the initialization period, the length of simulation runs, and the number of replications to be made of each run.

9. production runs and Analysis.
* Used to estimate measures of performances for the system designs that are being simulated.

10. More Runs?
* Based on the analysis of runs that have completed, the analyst determines if additional runs are needed and what designs those additional experiments should follow.

11. **Documentation and Reporting**

There are two types of documentation:

→ Program documentation: for the relationships between input parameters and output measures of performance and for a modification.

→ progress documentation: the history of a simulation a chronology of work done and decisions made.

12 Implementation: Implementation depends on how well

* Success of the implementation depends on how well the previous 11 steps have been performed.
* the model user has been thoroughly involved and
* If the model user understands the nature of the model and its outputs likelihood of a vigorous implementation is enhanced.

NOTE: The simulation model building can be broken into four phases.

Ist PHASE	IInd PHASE
* Step 1 and step 2 * It is a period of discovery and orientation * Analyst may have to restart the process if its not fine-tuned * Recalibrations & clarifications may occur in this phase or another phase	* Steps 3, 4, 5, 6, and 7 * A continuous interplay is required among the steps. * Exclusion of model user results in implications during implementation.
IIIrd PHASE	**IVth PHASE**
* steps 8, 9, and 10 * Conceives a thorough plan for experimenting * Discrete event stochastic is a statistical experiment * The output variables are estimated that contain random error and therefore proper statistical analysis is required.	* Steps 11 and 12 * Successful implementation depends on the involvement of user and every step's successful completion.

SIMULATION EXAMPLES

The simulations are carried out by the following three steps

1. Determine the characteristics of each of the inputs to the simulation.
   quite often, these are modeled as probability distributions, either continuous or discrete.

2. Construct a simulation table

Ex:-

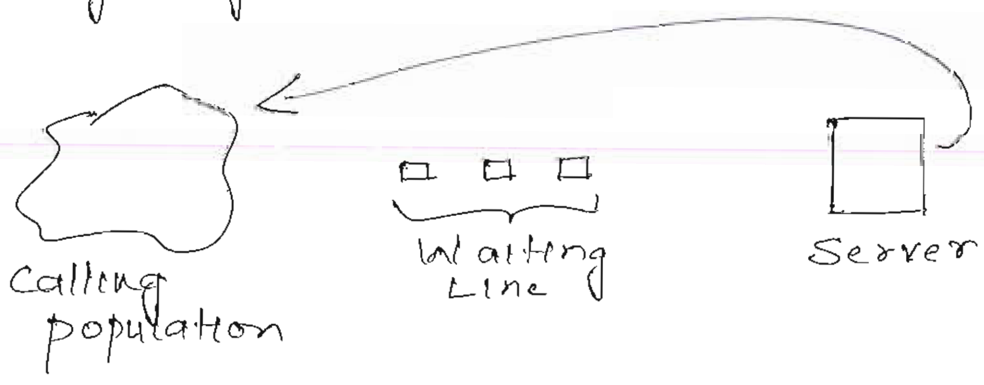
Repititions	Inputs				Response
	x_{i1}	x_{i2}	.....	x_{ip}	y_i
1					
2			....		
3					
⋮					
n					

3. For each repitition i, generate a value for each of the 'p' inputs, and evaluate a function, calculating a value of the response "y".

SIMULATION OF QUEUING SYSTEMS

* A queuing system is described by its calling population, the nature of arrivals, the service mechanism, the system capacity, and the queuing discipline.

• Fig shows an example for single channel queuing system.



* Single channel queuing System

→ Here calling population is infinite. ie their is no change in the arrival rate of the units.

→ Arrivals for service occur one at a time in a random fashion

→ service times are of some random length according to a probability distribution which does not change over time.

→ System capacity has no limit, meaning that any number of units can wait in line.

→ Finally, units are served in the order of their arrival.

* Definitions.

1. System state: Number of units in the system and the status of the server (Busy or Idle)

2. Event: A set of ~~instantaneous~~ circumstances ~~that~~ cause an instantaneous change in the state of a system.

In single channel system

→ Arrival Event: Entry of a unit into the system

→ Departure Event: Completion of service on a unit.

3. Simulation clock: Used to track simulated time.

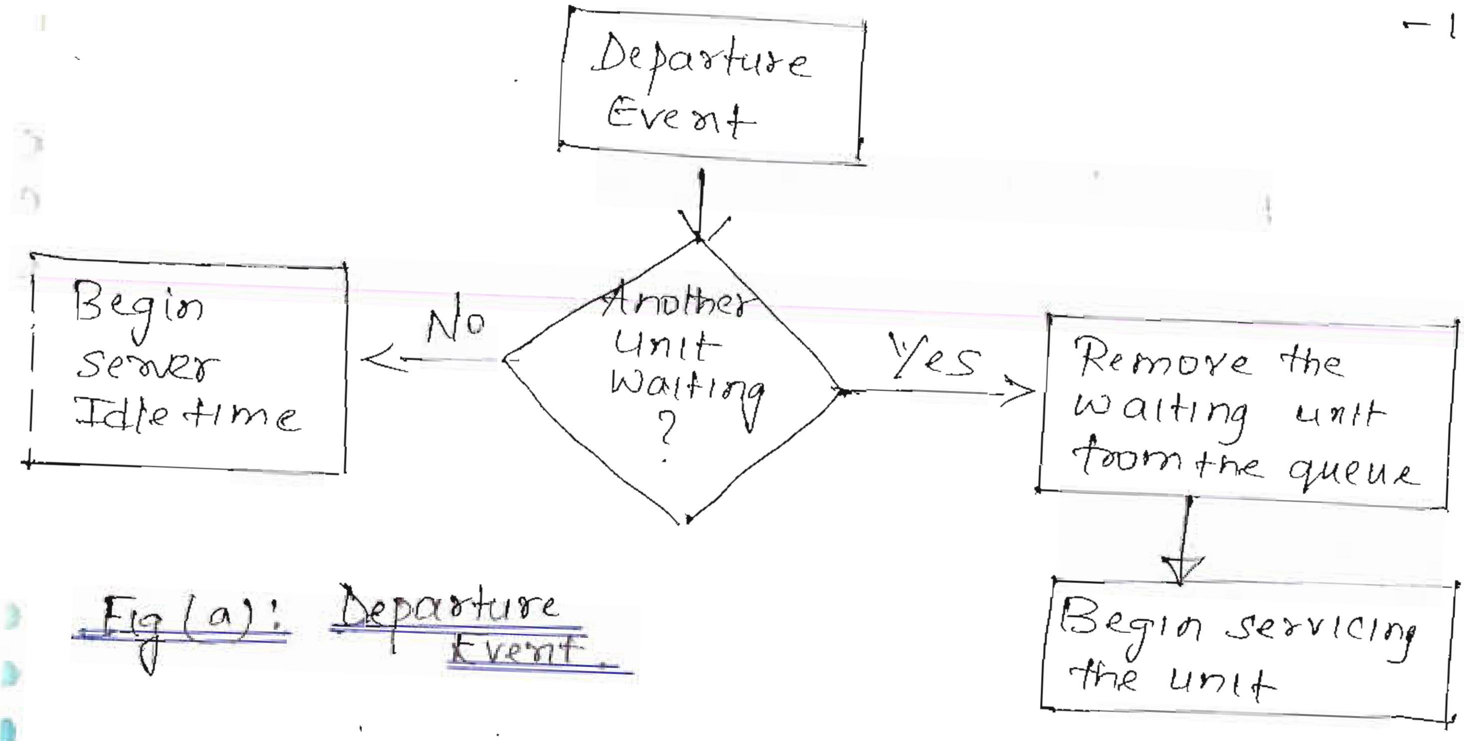


Fig (a): Departure Event

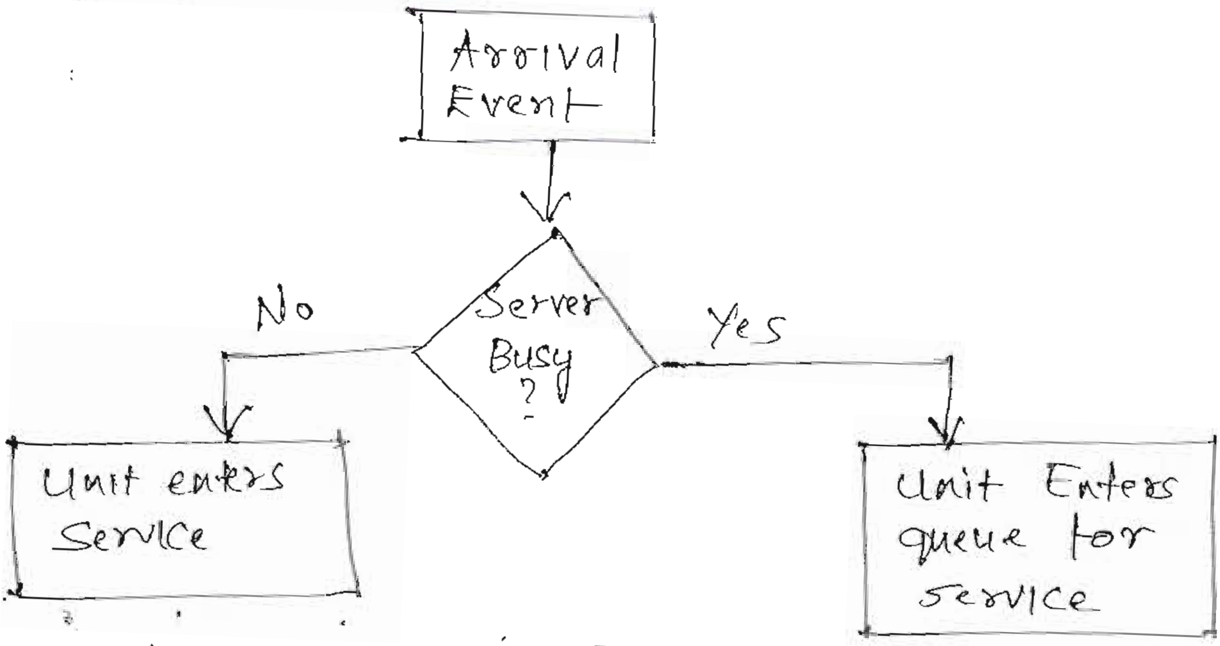


fig (b): Arrival Event

		Queue Status	
		Not Empty	Empty
Server status	Busy	Enter queue	Enter queue
	Idle	Impossible	Enter Service

Potential unit actions upon arrival

		Queue status	
		Not Empty	Empty
Server outcomes	Busy		Impossible
	Idle	Impossible	

Server outcomes after the completion of service

Single channel Queuing System.

problem 1: [May - June - 2010] - 10 Marks

A grocery store has one check out counter. customers arrives at this checkout counter at random from 1 to 8 minutes apart and each inter arrival time has the same probability of occurrence. The service time may vary from 1 to 6 minutes ~~apart~~, with probabilities given below.

Service (minutes)	1	2	3	4	5	6
probability	0.10	0.20	0.30	0.25	0.10	0.05

Simulate the arrival of 6 customers and calculate

→ Average waiting time for a customer
→ Probability that a customer has to wait
→ probability of a server being idle
→ Average Service time
→ Average time b/w Arrival.

Use the following sequence of Random Numbers

Random digits for Arrival	913	727	015	948	309	922
Random digits for service time	84	10	74	53	17	79

Assume that the first customer arrives at time 0. Depict the simulation in a tabular form.

Solution:

Here, Service time and Arrival time has not been given directly.
We have to externally find them out before constructing simulation table.
So the Solution consists of two steps.

Step 1: Calculating Service time and Arrival time from the given Random Numbers

Step 2: Constructing the Simulation table.

Step 1: Calculating Service time and Arrival time from the given Random Numbers.

Distribution of time b/w Arrivals.

Time b/w Arrivals (Minutes)	Probability	Cumulative probability	Random Digit Assignment.
1	0.125	0.125	001 – 125
2	0.125	0.250	126 – 250
3	0.125	0.375	251 – 375
4	0.125	0.500	376 – 500
5	0.125	0.625	501 – 625
6	0.125	0.750	626 – 750
7	0.125	0.875	751 – 875
8	0.125	1.000	816 – 000

Distribution of Service Time

Service Time minutes	P	C.P	R D A
1	0.10	0.10	01 – 10
2	0.20	0.30	11 – 30
3	0.30	0.60	31 – 60
4	0.25	0.85	61 – 85
5	0.10	0.95	86 – 95
6	0.05	1.00	96 – 00

Determining Time b/w Arrival

customer	R.D	Time b/w Arrival
1	–	–
2	913	8
3	727	6
4	015	1
5	948	8
6	309	3

Determing Service time

customer.	R.D	Service Time
1	84	4
2	10	1
3	74	4
4	53	3
5	17	2
6	79	4

Step 2 : Construction of Simulation Table

Customer	Inter arrival Time (min)	Arrival Time	Service Time (min)	Time Service Begins	Waiting Time in Queue (min)	Time Service Ends	Time Spent in System (min)	Idle time of server (min)
1	–	0	4	0	0	4	4	0
2	8	8	1	8	0	9	1	4
3	6	14	4	14	0	18	4	5
4	1	15	3	18	3	21	6	4
5	8	23	2	23	0	25	2	02
6	3	26	4	26	0	30	4	01
	Total :		18		3		21	16

1. Average Waiting Time $\Big\} = \dfrac{\text{Total time customer waits in Queue}}{\text{Total no. of customers}}$

$$= \frac{3}{6} = \boxed{0.5 \text{ minutes}}$$

2. probability that a customer has to wait in queue $\Big\} = \dfrac{\text{Total no. of Customers who wait}}{\text{Total no. of customers}}$

$$= \frac{1}{6} = \boxed{0.166}$$

$$= 16\%$$

3. probability of Idle time of server $\Big\} = \dfrac{\text{Total idle time of server}}{\text{Total run time of simulation}}$

$$= \frac{16}{30} = \boxed{0.533}$$

$$= 53.3\%$$

4. Average Service Time $\Big\} = \dfrac{\text{Total service time}}{\text{Total no. of customers}}$

$$= \frac{18}{6} = \boxed{3} \text{ minutes}$$

5. Average Time between Arrivals $= \dfrac{\text{Sum of all time b/w Arrivals}}{\text{number of arrivals} - 1}$

$= \dfrac{26}{5} = \boxed{5.2}$ minutes

6. Average waiting time of those who wait $= \dfrac{\text{Total time customers wait in Queue}}{\text{Total no. of customers } \cancel{\text{that}} \text{ wait}}$

$= \dfrac{3}{1} = \boxed{3}$ minutes

7. Average time customer spends in the system $= \dfrac{\text{Total time customers spend in the system}}{\text{Total no. of customers}}$

$= \dfrac{21}{6} = \boxed{3.5}$ minutes

Problem 2:

Consider the following continuously operating job shop. Using the following inter arrival and service time, construct a simulation table and perform simulation for 5 customers. Assume that when simulation begins, there is one job being processed.

a.) Find the average time in queue the customer spends

b.) What are the max. times the customer spends in system.

customer	1	2	3	4	5
Inter arrival time	–	0	60	60	120
Service time	25	50	37	45	50

Solution:

Customer	Inter Arrival Time	Arrival Time	Service Time	Time Service Begins	Waiting time in Queue	Time Service Ends	Time Customer spent in system	Idle time of server
1	–	0	25	0	0	25	25	0
2	0	0	50	25	25	75	75	0
3	60	60	37	75	15	112	52	0
4	60	120	45	120	0	165	45	8
5	120	240	50	240	0	290	50	75
					40			

a.) Average Time in queue a customer spends $\Big\} = \frac{40}{5} = \boxed{8}$ minutes

b.) Max time the customer spent in system $\Big\} = \boxed{75}$ minutes

Able Baker Car Hop problem

Here there is more than one service channel.

* Simulation proceeds in accordance with the following steps:

step 1: For Caller (or customer) k, generate an inter arrival time A_k. Add it to previous arrival time T_{k-1} to get the arrival time of the caller k.

step 2: If Able is Idle, Caller k begins service with Able at the current time T_{now}. If Able is Busy, but Baker is Idle, Caller k begins service with Baker at the current time T_{now}.

Step 3: If Able and Baker, both are busy, then service of Caller k, begins at the time at which the first one becomes available.

problem 3:

Cars arrive to the shop for service in the manner shown below

Time b/w Arrival	1	2	3	4
probability	0.25	0.40	0.20	0.15

There are two car hops - Able and Baker. Able works a Bit faster than Baker.
Perform the simulation for six callers and construct the simulation table for the same.

Distribution of their service time is shown below

For Able,

Service Time	2	3	4	5
P	0.30	0.28	0.25	0.17

For Baker,

Service Time	3	4	5	6
P	0.35	0.25	0.20	0.20

~~Solution~~:

Random digits for arrival are
26, 98, 90, 26, 42

Random digits for service are
95, 21, 51, 92, 89, 38.

Solution:

Step 1: Calculating Inter Arrival time and Service Time

Inter Arrival time Distribution

Time between Arrivals	P	C.P	R.D.A
1	0.25	0.25	01-25
2	0.40	0.65	26-65
3	0.20	0.85	66-85
4	0.15	1.00	86-00

Service Distribution

For Able,

Service Time	P	C.P	RDA
2	0.30	0.30	01-30
3	0.28	0.58	31-58
4	0.25	0.83	59-83
5	0.17	1.00	84-00

For Baker,

Service Time	P	C.P	RDA
3	0.35	0.35	0~~1~~ - 35
4	0.25	0.60	36 - 60
5	0.20	0.80	61 - 80
6	0.20	1.00	81 - 00

Inter Arrival Time Determination

Caller No.	Random Digit	Inter Arrival Time
1	-	-
2	26	2
3	98	4
4	90	4
5	26	2
6	42	2

Step 2: Construction of simulation table

Caller No.	Inter Arrival Time	Arrival Time	ABLE TSB	ABLE Service Time	ABLE TSE	BAKER TSB	BAKER Service Time	BAKER TSE	Caller Delay	Time In System
1	–	0	0	5	5	–	–	–	0	5
2	2	2	–	–	–	2	3	5	0	3
3	4	6	6	3	9	–	–	–	0	3
4	4	10	10	5	15	–	–	–	0	5
5	2	12	–	–	–	12	6	18	0	6
6	2	14	15	3	18	–	–	–	1	4

SIMULATION OF INVENTORY SYSTEM

* A simple inventory system is shown below:

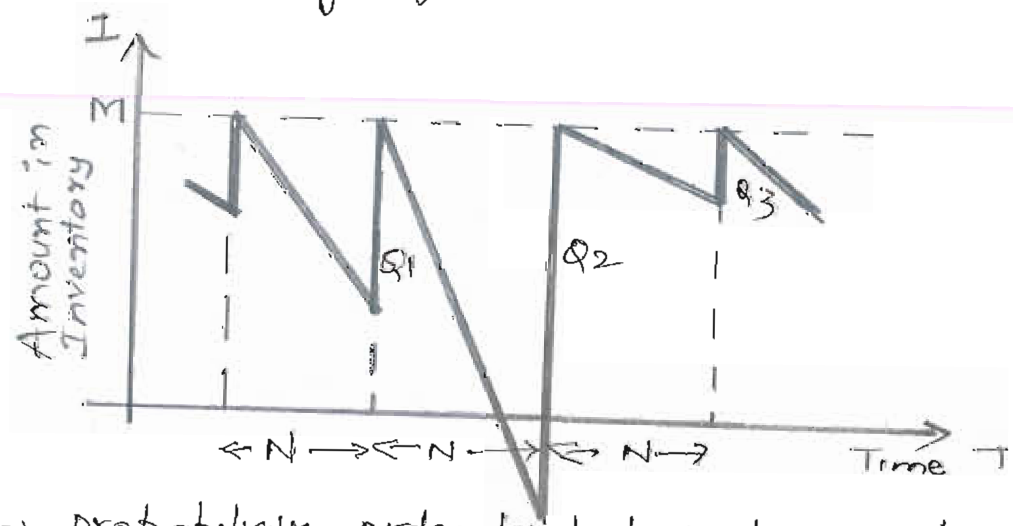


Fig: probabilistic order level inventory system.

* This inventory system has a periodic review of length N, at which time the inventory level is checked.

* An order is made to bring the inventory upto the level M. At the end of first review period, an order quantity Q1 is placed.

* In this inventory system, the lead time (length of time between the placement and receipt of an order) is zero.

* Notice that in the second cycle, the amount of in inventory drops below zero, indicating a shortage. In fig, these units are backordered. When the order arrives, the demand for backordered items are satisfied first. To avoid shortages, a buffer, or safety, stock would need to be carried.

* The total cost (or total profit) of an inventory system is the measure of performance.

In an (M, N) inventory system, the events that may occur are:

→ Demand for items in an inventory
→ The review of the inventory position.
→ Receipt of an order at the end of each review period.

When the lead time is zero (above fig), the last two events occurs simultaneously.

problem 4 :

The news stand buys the papers for 33 cents each and sells them for 50 cents each. Newspapers not sold at the end of the day are sold as scraps for 5 cents each. News papers can be purchased in bundles of 10. Thus news stand can buy 50, 60, 70 & so on. There are three types of days: 'Good', 'Fair', and 'poor'. They have the probabilities 0.35, 0.45, 0.20. Distribution of news papers demanded on each of these days is given below

Demand	Demand probability distribution		
	Good	Fair	Poor
40	0.03	0.10	0.44
50	0.05	0.18	0.22
60	0.15	0.40	0.16
70	0.20	0.20	0.12
80	0.35	0.08	0.06
90	0.15	0.04	0.00
100	0.07	0.00	0.00

Simulate the demand for 10 days and record profits from sales each day.
Find the optimal no. of papers the news stand should purchase.

~~Solution~~:

Given

Random Digits for types of news day

58, 17, 21, 45, 43, 36, 27, 73, 86, 19.

Random Digits for Demand

93, 63, 31, 19, 91, 75, 84, 37, 23, 02

Assume the news stand buys 70 newspapers each day.

Solution:

Step 1: Determining Type of news day and Demand.

Step 2: Construction of simulation table

Step 1: Determining Type of news day.

RDA for Type of news day

Type of News Day	P	C.P	R.D.A
Good	0.35	0.35	01 - 35
Fair	0.45	0.80	36 - 80
Poor	0.20	1.00	81 - 00

RDA for news papers demanded

Demand	P			C.P			RDA		
	Good	Fair	poor	Good	Fair	poor	Good	Fair	poor
40	0.03	0.10	0.44	0.03	0.10	0.44	01-03	01-10	01-44
50	0.05	0.18	0.22	0.08	0.28	0.66	04-08	11-28	45-61
60	0.15	0.40	0.16	0.23	0.68	0.82	09-23	29-68	67-82
70	0.20	0.20	0.12	0.43	0.88	0.94	24-43	69-88	83-94
80	0.35	0.08	0.06	0.78	0.96	1.00	44-78	89-96	95-00
90	0.15	0.04	0.00	0.93	1.00	1.00	79-93	97-00	←
100	0.07	0.00	0.00	1.00	1.00	1.00	94-00	←	←

Determining Type of News Day

Day	1	2	3	4	5	6	7	8	9	1
R.D	58	17	21	45	43	36	27	73	86	1
Type of News Day	Fair	Good	Good	Fair	Fair	fair	good	fair	poor	Good

Determining ~~type of~~ Demand

Day	1	2	3	4	5	6	7	8	9	10
R D	93	63	31	19	91	75	84	37	23	02
Demand	80	80	70	50	80	70	90	60	40	50

Step 2: Construction of Simulation Table

(profit) = [Revenue From sales] − [cost of news papers] − [lost profit from excess demand] + [Salvage from sale of scrap papers]

Day	Type of News Day	Demand	Revenue From Sales ($)	Lost profit from excess demand ($)	salvage from sales of scrap ($)	Daily profit ($)
1	Fair	80	35.00	1.7	—	10.20
2	Good	80	35.00	1.7	—	10.20
3	Good	70	35.00	—	—	11.90
4	Fair	50	25.00	—	1.00	2.90
5	Fair	80	35.00	1.7	—	10.20
6	Fair	70	35.00	—	—	11.90
7	Good	90	35.00	3.4	—	8.50
8	Fair	60	30.00	—	0.5	7.40
9	poor	40	20.00	—	1.5	−1.60
10	Good	40	20.00	—	1.5	−1.60
			$305.00	$8.5	$4.5	$70.00

Problem 5: June/July 2011 – 10 MARKS

Suppose the maximum inventory level is M, is 11 units and the review period, N is 5 days. Estimate by simulation, the average ending units in inventory and number of days when a shortage condition occurs.

The numbers of units demanded per day is given by the following probability distribution.
Assume that the orders are placed at the close of business and are received for inventory at the beginning of business as determined by the lead time.

Demand	0	1	2	3	4
probability	0.10	0.25	0.35	0.21	0.09

Initially the simulation has started with inventory level of 3 units and an order of 8 units scheduled to arrive in two-days time.
Lead time is a random variable, with the following probability distribution.

Lead time (Days)	1	2	3
probability	0.6	0.3	0.1

NOTE: The sequence of random digits for demand and random digits for lead time should be considered in the given order.

RD for Demand	24	35	65	81	54	3	87	27	73	70
	47	45	48	17	9					

RD for lead time	5	0	3

Solution:

RDA for Daily Demand

Demand	P	C.P	RDA
0	0.10	0.10	01 - 10
1	0.25	0.35	11 - 35
2	0.35	0.70	36 - 70
3	0.21	0.91	71 - 91
4	0.09	1.00	92 - 00

RDA for Lead time

Lead Time	P	CP	RDA
1	0.6	0.6	01 - 6
2	0.3	0.9	7 - 9
3	0.1	1.0	0

Formula:

$$\begin{pmatrix}\text{order}\\ \text{quantity}\end{pmatrix} = \begin{pmatrix}\text{order}\\ \text{up to}\\ \text{level}\end{pmatrix} - \begin{pmatrix}\text{Ending}\\ \text{Inventory}\end{pmatrix} - \begin{pmatrix}\text{shortage}\\ \text{quantity}\end{pmatrix}$$

Day	Cycle	Day within Cycle	Beginning Inventory	Random Digit for Demand	Demand	Ending Inventory	Shortage Quantity	Order Quantity	R.D. for lead time	Lead Time	Days until order arrives
1	1	1	3	24	1	2	0	—	—	—	1
2	1	2	2	35	1	1	0	—	—	—	—
3	1	3	8+1	65	2	7	0	—	—	—	—
4	1	4	7	81	3	4	0	—	—	—	—
5	1	5	3	54	2	1	0	10	5	1	1
6	2	1	1	3	0	1	0	—	—	—	—
7	2	2	11	87	3	8	0	—	—	—	—
8	2	3	8	27	1	7	0	—	—	—	—
9	2	4	7	73	3	4	0	—	—	—	—
10	2	5	4	70	2	2	0	9	0	3	3
11	3	1	2	47	2	0	0	—	—	—	2
12	3	2	0	45	2	0	2	—	—	—	1
13	3	3	0	48	2	0	4	—	—	—	—
14	3	4	9	17	1	4	0	—	—	—	—
15	3	5	4	9	0	4	0	7	3	1	1
					Total	45					

Thus,

Average Ending units in Inventory $= \frac{45}{15} = \boxed{3}$

Number of days when a shortage Condition occurs $= \boxed{2}$.

Problem 6: June/July 2011 - 10 MARKS

A large milling machine has three different bearings that fails in service. The c.d.f. of life of each bearing is identical as shown

Bearing in Life	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900
P	0.10	0.13	0.25	0.13	0.09	0.12	0.02	0.06	0.05	0.05

When the bearing fails, the mill stops, a repair person is called and a new bearing is installed. The delay time of repair person's arriving at the milling machine is also a random variable, with the distribution given below

Delay (minutes)	5	10	15
probability	0.6	0.3	0.1

Down time for the mill is estimated at \$ 5/min. The direct onsite cost of repair person is \$ 15/hour. It takes 20 min to change 1 bearing, 30 minutes to change 2 bearings, and 40 min to change 3 bearings. The bearing cost \$ 16 each. A proposal has been made to replace all 3 bearings whenever a bearing fails. Management needs an evaluation of this proposal. Simulate the system for 10,000 hours of operation under proposed ~~system~~ method and determine the total cost of proposed system.

Consider the following sequence of Random digits for bearing life times

Bearing 1	67	8	49	84	44	30	10	63
Bearing 2	70	43	86	93	81	44	19	51
Bearing 3	76	65	61	96	65	56	11	86

Consider the following sequence of Random Digits for delay time.

Delay	3	7	5	1	4	3	7	8

Solution:

Bearing Life Distribution

Bearing Life (Hours)	P	C.P	RDA
1000	0.10	0.10	01 - 10
1100	0.13	0.23	10 - 23
1200	0.25	0.48	24 - 48
1300	0.13	0.61	49 - 61
1400	0.09	0.70	62 - 70
1500	0.12	0.82	71 - 82
1600	0.02	0.84	83 - 84
1700	0.06	0.90	85 - 90
1800	0.05	0.95	91 - 95
1900	0.05	1.00	96 - 00

Delay Time Distribution

Delay Time (minutes)	P	CP	RDA
5	0.6	0.6	01 - 06
10	0.3	0.9	7 - 9
15	0.1	1.0	0

	Bearing 1		Bearing 2		Bearing 3		First Failure (Hours)	RP	Delay (minutes)
	R.D	Life (Hours)	RD	Life (Hours)	RD	Life (Hours)			
1	67	1400	70	1400	76	1500	1400	3	5
2	8	1000	43	1200	65	1400	1000	5	5
3	49	1300	86	1700	61	1300	1300	7	10
4	84	1600	93	1800	96	1900	1600	1	5
5	44	1200	81	1500	65	1400	1200	4	5
6	30	1200	44	1200	56	1300	1200	3	5
7	10	1100	19	1100	11	1100	1100	7	10
8	63	1400	51	1300	81	1700	1300	8	10
							Total: 10,100		55

Total cost of proposed system is estimated as follows.

$$\left.\begin{array}{l}\text{cost of}\\ \text{Bearings}\end{array}\right\} = (8 \times 3) * (\$32/\text{bearing})$$
$$= 24 * 32$$
$$= \underline{\underline{\$768}}$$

$$\left.\begin{array}{l}\text{Cost of}\\ \text{delay time}\end{array}\right\} = 55 \text{ min} * \$10/\text{min}$$
$$= \underline{\underline{\$550}}$$

$$\left.\begin{array}{l}\text{cost of downtime}\\ \text{during repair}\end{array}\right) = 8\text{sets} * 40 \text{ min/set} * \$5/\text{min}$$
$$= \underline{\underline{\$1600}}$$

$$\left.\begin{array}{l}\text{Cost of repair}\\ \text{person}\end{array}\right\} = 8\text{sets} * 40\text{min/set} * \$30/60\text{ min}$$
$$= \frac{8 * 40 * 30}{60}$$
$$= \underline{\underline{\$160}}$$

$$\left(\begin{array}{c}\text{Total}\\ \text{cost}\end{array}\right) = \$768 + \$550 + \$1600 + \$160$$
$$= \underline{\underline{\$3078}}$$

Total life of Bearings is 10,100 * 3 = 30,300 hours

Therefore,

$$\left.\begin{array}{l}\text{Total cost per}\\ \text{10,000 bearings}\\ \text{hours}\end{array}\right\} = \frac{\$3078}{3.03} = \boxed{\$\,1015.841}$$

UNIT 2:

GENERAL PRINCIPLES, SIMULATION SOFTWARE

CONCEPTS IN DISCRETE-EVENT SIMULATION

- **System:** A collection of entities (e.g., people and machines) that interact together over time to accomplish one or more goals.
- **Model:** An abstract representation of a system, usually containing structural, logical, or mathematical relationships which describe a system in terms of state, entities and their attributes, sets, processes, events, activities, and delays.
- **System state:** A collection of variables that contain all the information necessary to describe the system at any time.
- **Entity:** Any object or component in the system which requires explicit representation in the model (e.g., a server, a customer, a machine).
- **Attributes:** The properties of a given entity (e.g., the priority of a v customer, the routing of a job through a job shop).
- **List:** A collection of (permanently or temporarily) associated entities ordered in some logical fashion (such as all customers currently in a waiting line, ordered by first come, first served, or by priority).
- **Event:** An instantaneous occurrence that changes the state of a system as an arrival of a new customer).
- **Event notice:** A record of an event to occur at the current or some future time, along with any associated data necessary to execute the event; at a minimum, the record includes the event type and the event time.
- **Event list:** A list of event notices for future events, ordered by time of occurrence; also known as the future event list (FEL).
- **Activity:** A duration of time of specified length (e.g., a service time or arrival time), which is known when it begins (although it may be defined in terms of a statistical distribution).
- **Delay:** A duration of time of unspecified indefinite length, which is not known until it ends (e.g., a customer's delay in a last-in, first-out waiting line which, when it begins, depends on future arrivals).
- **Clock:** A variable representing simulated time.

EXAMPLE (Able and Baker, Revisited)
Consider the Able-Baker carhop system of explained in unit 1. A discrete- event model has the following components:

System state
L_Q (t), the number of cars waiting to be served at time t
$L_A(t)$, 0 or 1 to indicate Able being idle or busy at time t
$L_B(t)$, 0 or 1 to indicate Baker being idle or busy at time t

ASHOK KUMAR K
Mob: 9742024066
e-mail: celestialcluster@gmail.com

Entities
Neither the customers (i.e., cars) nor the servers need to be explicitly represented, except in terms of the state variables, unless certain customer averages are desired
Events
Arrival event
Service completion by Able
Service completion by Baker
Activities
Interarrival time
Service time by Able
Service time by Baker
Delay
A customer's wait in queue until Able or Baker becomes free.

Event Scheduling/ Time Advance Algorithm

- After the system snapshot at simulation time CLOCK = t has been updated, the CLOCK is advanced to simulation time CLOCK = t1 the imminent event notice is removed from the FEL, and the event is executed
- At time t1, new future events may or might not be generated, but if any are, they are scheduled by creating event notices and putting them in to their proper position on the FEL.
- After the new system snapshot for time t1 has been updated, the clock is advanced to the time of the new imminent event and that event is executed.
- This process repeats until the simulation is over
- The sequence of actions that a simulator must perform to advance the clock and build a new system snapshot is called the event-scheduling/time-advance algorithm
- The removal and addition of events from FEL is illustrated in below figure:

ASHOK KUMAR K
Mob: 9742024066
e-mail: celestialcluster@gmail.com

Old system snapshot at time t

CLOCK	System State	...	Future Event List	...
t	(5, 1, 6)		$(3, t_1)$ — Type 3 event to occur at time t_1 $(1, t_2)$ — Type 1 event to occur at time t_2 $(1, t_3)$ — Type 1 event to occur at time t_3 . . . $(2, t_n)$ — Type 2 event to occur at time t_n	

Event-scheduling/time-advance algorithm

Step 1. Remove the event notice for the imminent event (event 3, time t_1) from FEL

Step 2. Advance CLOCK to imminent event time (i.e., advance CLOCK from t to t_1).

Step 3. Execute imminent event: update system state, change entity attributes, and set membership as needed.

Step 4. Generate future events (if necessary) and place their event notices on FEL ranked by event time. (*Example:* Event 4 to occur at time t^*, where $t_2 < t^* < t_3$.)

Step 5. Update cumulative statistics and counters.

New system snapshot at time t_1

CLOCK	System State	...	Future Event List	...
t_1	(5, 1, 5)		$(1, t_2)$ — Type 1 event to occur at time t_2 $(4, t^*)$ — Type 4 event to occur at time t^* $(1, t_3)$ — Type 1 event to occur at time t_3 . . . $(2, t_n)$ — Type 2 event to occur at time t_n	

- Event 3 with event time t_1 represents, say, a service completion event at server 3. Since it is the imminent event at time t, it is removed from the FEL in step 1 of the event scheduling/time advance algorithm. When event 4 (say, an arrival event) with event time t^* is generated at step 4, one possible way to determine its correct position on the FEL is to conduct a top-down search:

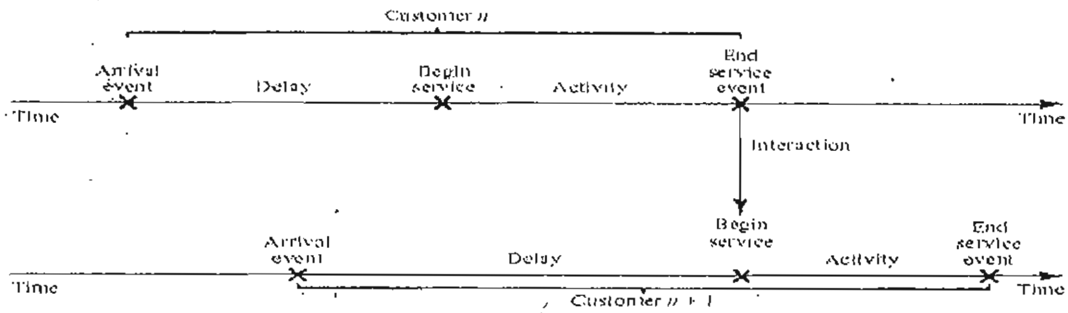
If $t^* < t_2$,	place event 4 at the top of the FEL.
If $t_2 < t^* < t_3$,	place event 4 second on the list.
If $t_3 < t^* < t_4$,	place event 4 third on the list.
	. . .
If $t_n < t^*$,	event 4 last on the list.

World views

- When using a simulation package or even when using a manual simulation, a modeler adopts a **world view** or orientation for developing a model.
- The most prevalent world views are:
    - Event-scheduling world view
    - Process-interaction world view
    - Activity-scanning world view

ASHOK KUMAR K
Mob: 9742024066
e-mail: celestialcluster@gmail.com

- **Event-scheduling world view:**
  - When using the event-scheduling approach, a simulation analyst concentrates on events and their effect on system state.
- **Process-interaction world view:**
  - When using the process-interaction approach, a simulation analyst thinks in terms of processes.
  - The analyst defines the simulation model in terms of entities or objects and their life cycle as they flow through the system, demanding resources and queuing to wait for resources.
  - More precisely, a process is the life cycle of one entity. This life cycle consists of various events and activities.
  - Figure shows the interaction between two customer processes as customer n+1 is delayed until the previous customer's "end-service event" occurs



- **Activity-scanning world view:**
  - It uses a fixed time increment and a rule-based approach to decide whether any activities can begin at each point in simulated time.
  - With this approach, the modeler concentrates on the activities of a model and those conditions that allow an activity to begin.
  - In the three-phase approach, activities are divided into two categories.
    - **B activities:** activities bound to occur; all primary events and unconditional activities.
    - **C activities:** activities or events that are conditional upon certain conditions being true.
  - With the three phase approach, the simulation proceeds with repeated execution of the 3 phases until it is completed.
    - **Phase A:** Remove the imminent event from the FEL and advance the clock to its event time. Remove any other events from the FEL that have the same event time.
    - **Phase B:** Execute all B-type events that were removed from the FEL.
    - **Phase C:** Scan the conditions that trigger each C-type activity and activate any whose conditions are met. Rescan until no additional C-type activities can begin or events occur.

ASHOK KUMAR K
Mob: 9742024066
e-mail: celestialcluster@gmail.com

Example : Single channel Queue.

Given

Inter arrival times	1	1	6	3	7	5	2	4	1 ..
service times	4	2	5	4	1	5	4	1	4 ..

~~perform~~ simulation until sixth customer departs.
Find server utilization and maximum queue length

Soln:

* Server utilization $= \dfrac{\text{Total server busy time } (B)}{\text{total time } (T_E)}$

* Initial conditions:
  - → $L_Q(0) = 0$
  - → $L_S(0) = 1$
  - → Write both departure & Arrival event on FEL.

* Here $a^* =$ generated interarrival time
  $s^* =$ generated service time.

	System State				Cumulative Statistics	
Clock	$LQ(t)$	$LS(t)$	Future Event List	Comment	B	MQ
0	0	1	(A, 1) (D, 4) (E, 60)	First A occurs ($a^* = 1$) Schedule next A ($s^* = 4$) Schedule first D	0	0
1	1	1	(A, 2) (D, 4) (E, 60)	Second A occurs: (A, 1) ($a^* = 1$) Schedule next A (Customer delayed)	1	1
2	2	1	(D, 4) (A, 8) (E, 60)	Third A occurs: (A, 2) ($a^* = 6$) Schedule next A (Two customers delayed)	2	2
4	1	1	(D, 6) (A, 8) (E, 60)	First D occurs: (D, 4) ($s^* = 2$) Schedule next D (Customer delayed)	4	2
6	0	1	(A, 8) (D, 11) (E, 60)	Second D occurs: (D, 6) ($s^* = 5$) Schedule next D	6	2

(continued overleaf)

	System State				Cumulative Statistics	
Clock	$LQ(t)$	$LS(t)$	Future Event List	Comment	B	MQ
8	1	1	(D, 11) (A, 11) (E, 60)	Fourth A occurs: (A, 8) ($a^* = 3$ Schedule next A (Customer delayed)	8	2
11	1	1	(D, 15) (A, 18) (E, 60)	Fifth A occurs: (A, 11) ($a^* = 7$) Schedule next A Third D occurs: (D, 11) ($s^* = 4$) Schedule next D Customer delayed	11	2
15	0	1	(D, 16) (A, 18) (E, 60)	Fourth D occurs: (D, 15) ($s^* = 1$) Schedule next D	15	2
16	0	0	(A, 18) (E, 60)	Fifth D occurs: (D, 16)	16	2
18	0	1	(D, 23) (A, 23) (E, 60)	Sixth A occurs ($a^* = 5$) Schedule next A ($s^* = 5$) Schedule next D	16	2
23	0	1	(A, 25) (D, 27) (E, 60)	Seventh A occurs: (A, 23) ($a^* = 2$) Schedule next Arrival Sixth D occurs: (D, 23)	21	2

Ex: checkout counter continued.

Estimate the mean response time and mean proportion of customers who spend 5 or more minutes in the system.

Sol(n):

* Response time is the length of time a customer spends in the system.

  Response time = CLOCKTIME − attribute "time of arrival"

* Here there are some additional model components.

  Entities: (C_i, t) representing customer C_i who arrived at time t

  Event notices:

  (A, t, C_i): Arrival of customer C_i at future time t

  (D, t, C_j): Departure of customer C_j at future time t.

  Set: "CHECKOUT LINE", the set of all customers currently at the checkout counter (being served or waiting to be served), ordered by time of arrival.

* Three new cumulative statistics will be collected

  $S \rightarrow$ sum of customer response times for all customers who have departed by the current time.

  $F \rightarrow$ total no. of customers who spend 5 or more minutes at the check out counter.

  $N_D \rightarrow$ total no. of departures up to the current simulation time

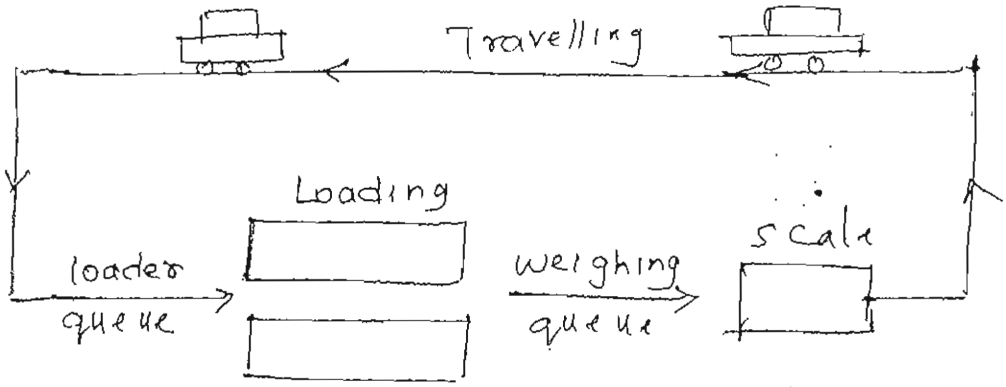
Table 3.2 Simulation Table for Example 3.4

System State					Cumulative Statistics		
Clock	$LQ(t)$	$LS(t)$	CHECKOUT LINE	Future Event List	S	N_D	F
0	0	1	(C1,0)	(A,1,C2) (D,4, C1) (E,60)	0	0	0
1	1	1	(C1,0) (C2,1)	(A,2,C3) (D,4,C1) (E,60)	0	0	0
2	2	1	(C1,0) (C2,1) (C3,2)	(D,4,C1) (A,8,C4) (E,60)	0	0	0
4	1	1	(C2,1) (C3,2)	(D,6,C2) (A,8,C4) (E, 60)	4	1	0
6	0	1	(C3,2)	(A,8,C4) (D,11,C3) (E,60)	9	2	1
8	1	1	(C3,2) (C4,8)	(D,11,C3) (A,11,C5) (E,60)	9	2	1
11	1	1	(C4,8) (C5,11)	(D,15,C4) (A,18,C6) (E,60)	18	3	2
15	0	1	(C5,11)	(D,16,C5) (A,18,C6) (E,60)	25	4	3
16	0	0		(A,18,C6) (E,60)	30	5	4
18	0	1	(C6,18)	(D,23,C6) (A,23,C7) (E,60)	30	5	4
23	0	1	(C7,23)	(A,25,C8) (D,27,C7) (E,60)	35	6	5

For a simulation run length of 23 min, the average response time was $\frac{S}{N_D} = \frac{35}{6} = \boxed{5.83 \text{ min}}$

and the observed proportion of customers who spent 5 or more minutes in the system was $\frac{F}{N_D} = \boxed{0.83}$.

The Dump-Truck problem.



* The purpose of this simulation is to estimate the loader & scale utilizations (percentage of time busy)
* The model has the following components:

System state $[LQ(t), L(t), WQ(t), W(t)]$

$LQ(t)$ = no. of trucks in loader queue
$L(t)$ = no. of trucks (0, 1, or 2) being loaded
~~L~~. $WQ(t)$ = no. of trucks in weigh queue
$W(t)$ = no. of trucks (0, or 1) being weighed, all at simulation time t.

Event notices

(ALQ, t, DT_i) : dump truck i arrives at Loader queue (ALQ) at time t.

(EL, t, DT_i) : Dump truck i ends loading (EL) at time t.

(EW, t, DT_i) : Dump truck i ends weighing (EW) at ~~loader~~ time t.

Entities

six dump trucks (DT_1 ... DT_6) ⟨Assumption⟩

Lists

Loader queue, all trucks ~~weight~~ waiting to begin loading, ordered on FCFS basis.

weigh queue, all trucks waiting to be weighed, ordered on FCFS basis

Activities

Loading time, weighing time, travel time.

Delays

Delay at loader queue, & delay at scale.

problem:

Six dump trucks are used to haul coal from the entrance of a mine to railroad. Each truck is loaded by one of two loaders. After loading, a truck immediately moves to the scale, to be weighed as soon as possible. Both the loaders & the scale have a FCFS waiting line for trucks. Travel time from a loader to scale is considered negligible. After being weighed, a truck begins travel time (during which time truck unloads) and then afterwards returns to the loader queue. The activities of loading time, weighing time, & travel time are given below.

loading time	10	5	5	10	15	10	10
weighing time	12	12	12	16	12	16	
travel time	60	100	40	40	80		

(End of simulation is completion of two weighings from the scale) X

Depict the simulation table and estimate the loader and scale utilizations.

Assume that five of the trucks are at the loaders and one is at the scale at time 0

– (July 08, 10 marks).

– (Feb 06, 15 marks)

stopping time is $T_E = 76$ min.

24089

In order to estimate loader & scale utilizations, two cumulative statistics are maintained:

B_L: total busy time of both loaders from time 0 to time t

B_S: total busy time of the scale from time 0 to time t.

Clock t	System State $LQ(t)$	$L(t)$	$WQ(t)$	$W(t)$	Lists: Loader Queue	Lists: Weigh Queue	Future Event List	Cumulative Statistics B_L	B_S
0	3	2	0	1	DT4 DT5 DT6		(EL, 5, DT3) (EL, 10, DT2) (EW, 12, DT1)	0	0
5	2	2	1	1	DT5 DT6	DT3	(EL, 10, DT2) (EL, 5 + 5, DT4) (EW, 12, DT1)	10	5
10	1	2	2	1	DT6	DT3 DT2	(EL, 10, DT4) (EW, 12, DT1) (EL, 10 + 10, DT5)	20	10
10	0	2	3	1		DT3 DT2 DT4	(EW, 12, DT1) (EL, 20, DT5) (EL, 10 + 15, DT6)	20	10
12	0	2	2	1		DT2 DT4	(EL, 20, DT5) (EW, 12 + 12, DT3) (EL, 25, DT6) (ALQ, 12 + 60, DT1)	24	12
20	0	1	3	1		DT2 DT4 DT5	(EW, 24, DT3) (EL, 25, DT6) (ALQ, 72, DT1)	40	20
24	0	1	2	1		DT4 DT5	(EL, 25, DT6) (EW, 24 + 12, DT2) (ALQ, 72, DT1) (ALQ, 24 + 100, DT3)	44	24
25	0	0	3	1		DT4 DT5 DT6	(EW, 36, DT2) (ALQ, 72, DT1) (ALQ, 124, DT3)	45	25
36	0	0	2	1		DT5 DT6	(EW, 36 + 16, DT4) (ALQ, 72, DT1) (ALQ, 36 + 40, DT2) (ALQ, 124, DT3)	45	36
52	0	0	1	1		DT6	(EW, 52 + 12, DT5) (ALQ, 72, DT1) (ALQ, 76, DT2) (ALQ, 52 + 40, DT4) (ALQ, 124, DT3)	45	52
64	0	0	0	1			(ALQ, 72, DT1) (ALQ, 76, DT2) (EW, 64 + 16, DT6) (ALQ, 92, DT4) (ALQ, 124, DT3) (ALQ, 64 + 80, DT5)	45	64

Clock t	System State $LQ(t)$	$L(t)$	$WQ(t)$	$W(t)$	Lists Loader Queue	Weigh Queue	Future Event List	Cumulative Statistics B_L	B_S
72	0	1	0	1			(ALQ, 76, DT2) (EW, 80, DT6) (EL, 72 + 10, DT1) (ALQ, 92, DT4) (ALQ, 124, DT3) (ALQ, 144, DT5)	45	[illegible]
76	0	2	0	1			(EW, 80, DT6) (EL, 82, DT1) (EL, 76 + 10, DT2) (ALQ, 92, DT4) (ALQ, 124, DT3) (ALQ, 144, DT5)	49	[illegible]

$$\left.\begin{array}{l}\text{Average loader}\\ \text{utilization}\end{array}\right\} = \frac{49/2}{76} = \boxed{0.32}$$

$$\left.\begin{array}{l}\text{Average scale}\\ \text{utilization}\end{array}\right\} = \frac{76}{76} = \boxed{1.00}$$

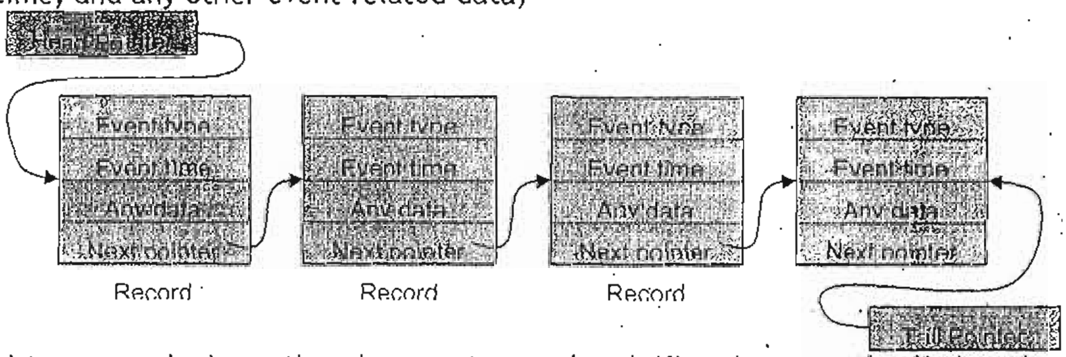
2224310

LIST PROCESSING

- **List processing** deals with methods for handling lists of entities and the future event list.

Lists: Basic properties and operations

- **List:** a set of ordered or ranked records.
- **Record:** one entity or one event notice.
- **Field** : an entity identifier and its attributes (it may be the event type, event time, and any other event related data)



- Lists are ranked, so they have a top or head (first item on the list) and a bottom or tail (the last item on the list)
- The main operations on a list are:
  - **Removing a record from the top of the list.**
    - When time is advanced and the imminent event is due to be executed.
    - By adjusting the head pointer on the FEL ⇔ by removing the event at the top of the FEL.
  - **Removing a record from any location on the list.**
    - If an arbitrary event is being canceled, or an entity is removed from a list based on some of its attributes (say, for example, its priority and due date) to begin an activity.
    - By making a partial search through the list.
  - **Adding an entity record to the top or bottom of the list.**
    - When an entity joins the back of a first-in first-out queue.
    - By adjusting the tail pointer on the FEL ⇔ by adding an entity to the bottom of the FEL
  - **Adding a record to an arbitrary position on the list, determined by the ranking rule.**
    - If a queue has a ranking rule of earliest due date first (EDF).
    - By making a partial search through the list.

ASHOK KUMAR K
Mob: 9742024066
e-mail: celestialcluster@gmail.com

Using arrays for list processing

- The notation R(i): the i^{th} record in the array
- **Advantage:** Any specified record, say the ith, can be retrieved quickly without searching, merely by referencing R(i).
- **Disadvantages:**
  - When items are added to the middle of a list or the list must be rearranged.
  - Arrays typically have a fixed size, determined at compile time or upon initial allocation when a program first begins to execute.
  - In simulation, the maximum number of records for any list may be difficult or impossible to determine ahead of time, while the current number in a list may vary widely over the course of the simulation run.
- In the use of arrays for sorting lists, there are two basic methods for keeping track of the ranking of records in a list
  - To store the first record in R(1), the second in R(2), and so on, and the last in R(tailptr), where tailptr is used to refer to the last item in the list.
  - A variable called a head pointer, with name *headptr*, points to the record at the top of the list.

SIMULATION IN JAVA

- Java is a widely used programming language that has been used extensively in simulation.
- It does not provide any facilities directly, so the simulation analyst must program all details of event-scheduling/time-advance algorithm, the statistic-gathering capability etc. however the runtime library does provide a random number generator.
- The following components are common to almost all models written in java:
  - **Clock:** A variable defining simulated time.
  - **Initialization method:** A method to define the system state at time 0.
  - **Min-time event method:** A method that identifies the imminent event. (Future Event List) that has the smallest time stamp.
  - **Event methods:** for each event type, a method to update system state when that event occurs.
  - **Random-variate generators:** Methods to generate samples from desired probability distribution.
  - **Main Program:** To maintain overall control of the event scheduling algorithm.
  - **Report generator:** Computes summary from cumulative statistics and prints a report at the end of simulation.

ASHOK KUMAR K
Mob: 9742024066
e-mail: celestialcluster@gmail.com

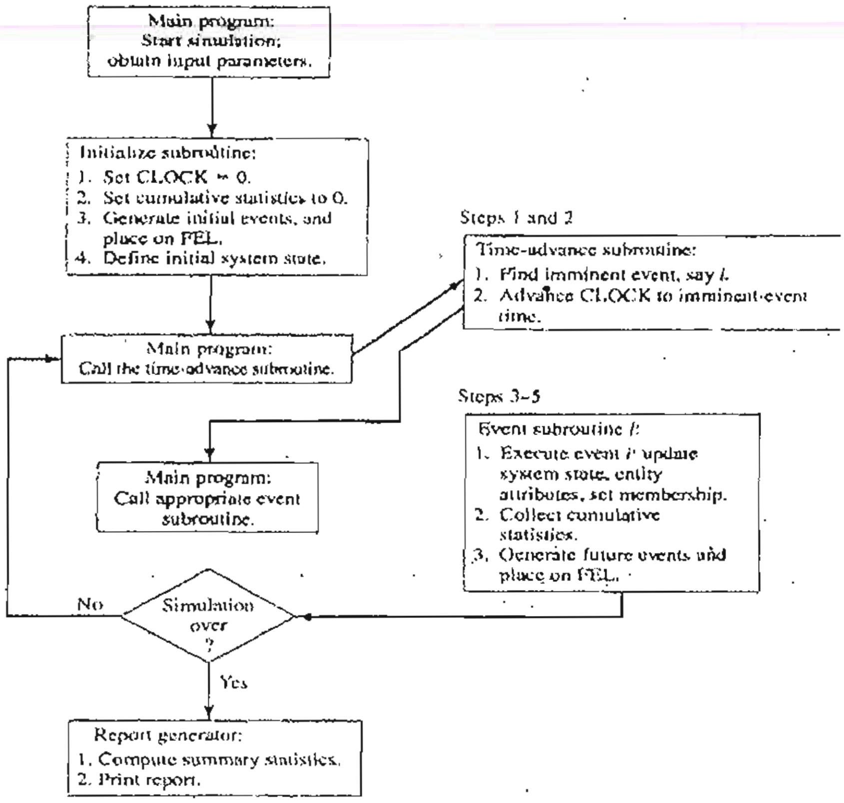


Fig: overall structure of an event-scheduling simulation program

ASHOK KUMAR K
Mob: 9742024066
e-mail: celestialcluster@gmail.com

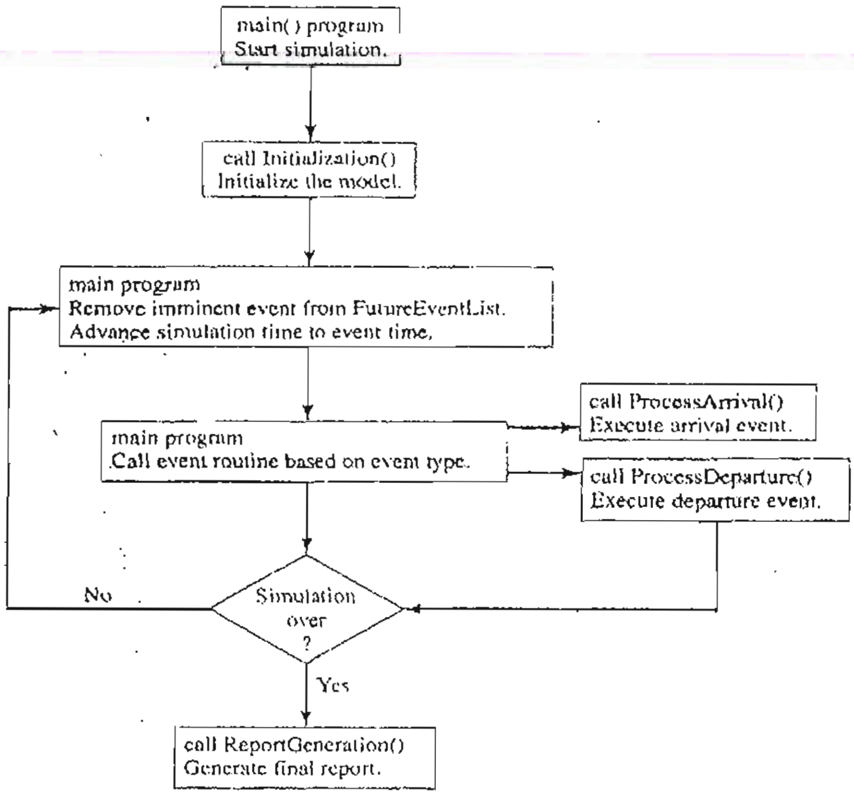


Fig: overall structure of java simulation of a single-server queue.

SIMULATION IN GPSS

- GPSS is a highly structured, special purpose simulation programming language based on the process-interaction approach and oriented toward queuing systems.
- There are over 40 standard blocks in GPSS. Entities called transactions may be viewed as flowing through the block diagram. Blocks represent events, delays, and other actions that affect transaction flow.
- Thus, GPSS can be used to model any situation where transactions (entities, customers, units of traffic) are flowing through a system. The block diagram is converted into block statements, control statements are added, and the result is a GPSS model.

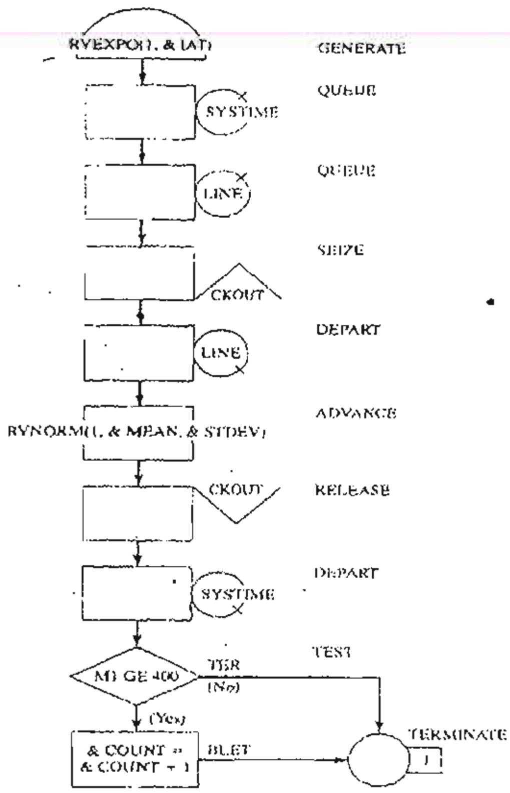


Fig: GPSS block diagram for the single-server queue simulation

ASHOK KUMAR K
Mob: 9742024066
e-mail: celestialcluster@gmail.com

UNIT 3 :

STATISTICAL MODELS IN SIMULATION

Syllabus

* Review of terminology and concepts.
* Useful statistical models
* Discrete Distributions
* Continuous Distributions
* poissons process
* Empirical Distributions

— 6 Hours.

Ashok KUMAR. K.
9742024066
celestialcluster@gmail.com

* To model the real world phenamena, we have to represent the actions of the entities within the system. These actions are usually <u>unpredictable</u>.
* It turns out that some <u>statistical models</u> are suitable to describe such actions.
* To model a random process (eg: service time) the modeler has to select an appropriate <u>probability distribution</u>.
* steps to select an appropriate p.d
  → Sample the process
  → Select the known distribution (experience, special software)
  → Estimate the values of the parameters of the distribution.
  → Test to see how good is the fit

REVIEW OF TERMINOLOGY AND CONCEPTS

→ Discrete Random Variables
→ Continuous Random Variables
→ Cumulative Distribution Function
→ Expectation
→ Mode.

Discrete Random Variables

* Random Variable:
A variable that assumes values in a irregular pattern is called a Random Variable.

* Discrete Random Variables:

Let X be a random variable. If the number of possible values of X is finite or countably infinite, then X is called a descrete random variable.

Let x_i be all possible values of X, and $p(x_i) = P(X = x_i)$ be the probability that $X = x_i$ then $p(x_i)$ must meet the following conditions

$$\boxed{\begin{array}{l} 1. \quad p(x_i) \geq 0 \quad \text{for all } i. \\ 2. \quad \sum_{i=1}^{\infty} p(x_i) = 1 \end{array}}$$

* The collection of pairs $(x_i, P(x_i))$ $i = 1, 2, \ldots$ is called the probability distribution of X, and $p(x_i)$ is called probability mass function (pmf) of X.

Continuous Random Variables

* if the values of X is an interval or a collection of intervals, then X is called a <u>contineous Random Variable</u>.

* For a contineous Random variable, its probability is represented by:-

$$P(a \le X \le b) = \int_a^b f(x)\,dx \quad \text{——— (1)}$$

the function $f(x)$ is called the <u>probability density function</u> (p.d.f) of X, which has to meet the following conditions

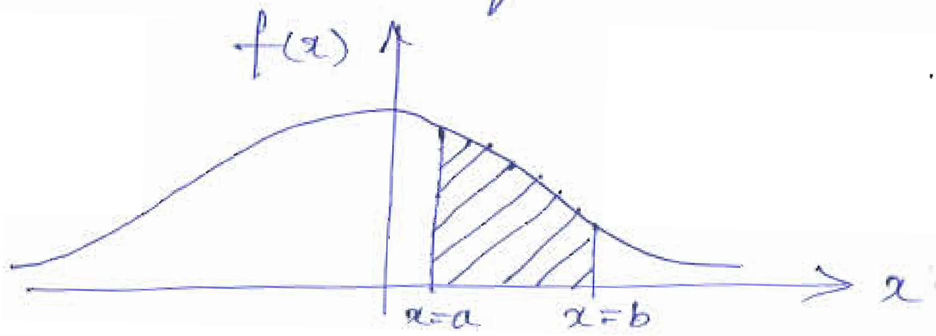
1. $f(x) \ge 0$ for all x in R_X (Range space)
2. $\int_{R_X} f(x)\,dx = 1$
3. $f(x) = 0$ if x is not in R_X

* <u>Note</u>,

$P(a \le x \le b) = P(a < x \le b) = P(a \le x < b) = P(a < x < b)$

$$\left[\because \int_{x_0}^{x_0} f(x)\,dx = 1 \right]$$

* Graphical Representation of eq(1) is shown below.



shaded area represents the probability that X lies in the interval (a,b)

Problem 1:

Life of an inspection device is given by X, a contineous Random variable with p.d.f :-

$$f(x) = \begin{cases} \frac{1}{2} e^{-x/2} & , x \geq 0 \\ 0 & , \text{otherwise} \end{cases}$$

Find the probability that the device's life cycle is b/w 2 and 3 years.

Solution:

$a = 2$

$b = 3$

$$P(2 < X < 3) = \int_2^3 f(x)\,dx$$

$$= \int_2^3 \frac{1}{2} e^{-x/2}\,dx$$

$$= \frac{1}{2}\left(\frac{e^{-x/2}}{-1/2}\right)_2^3$$

$$= \frac{1}{2}\left[\frac{e^{-3/2}}{-1/2} + \frac{e}{1/2}\right]$$

$$= -0.223 + 0.368$$

$$\therefore \boxed{P(2 < X < 3) = 0.145}$$

Cumulative Distribution Function (cdf)

* The cumulative distribution function (c.d.f) denoted by $F(x)$ measures the probability that the random variable X assumes the value less than or equal to x.

i.e $F(x) = P(X \leq x)$

→ If X is discrete,

$$F(x) = \sum_{\substack{all \\ x_i \leq x}} P(x_i)$$

→ If X is contineous,

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$

* properties of C.D.F,

1. F is a non decreasing function. if $(a<b)$. then $F(a) < F(b)$.
2. $\lim_{x \to \infty} F(x) = 1$
3. $\lim_{x \to -\infty} F(x) = 0$.

* All probability questions about X can be answered in terms of c.d.f

Ex: $P(a<X<b) = F(b) - F(a) \ \forall \ a<b$.

Problem 2:

An inspection device has c.d.f

$$F(x) = \frac{1}{2}\int_0^x e^{-t/2}\,dt.$$

Find

a.) probability that the device lasts for less than 2 years.

b.) probability that the device lasts between 2 and 3 years.

Solution:

a.) $$P(0 \le X \le 2) = F(2) - F(0)$$
$$= F(2)$$
$$= \frac{1}{2}\int_0^2 e^{-t/2}\,dt$$
$$= \frac{1}{2}\left(\frac{e^{-t/2}}{1/2}\right)_0^2$$

$$\boxed{P(0 < x < 2) = 0.632}$$

b.) $$P(2 < x < 3) = F(3) - F(2)$$
$$= \frac{1}{2}\int_0^3 e^{-t/2}\,dt - 0.632$$
$$= \frac{1}{2}\left(\frac{e^{-t/2}}{1/2}\right)_0^3 - 0.632$$

$$\boxed{P(2 < x < 3) = 0.145}$$

Expectation and Variance.

* Expectation essentially is the expected value of a random Variable.
* Variance is a measure of how a random variable varies from its expected value.

* For Discrete random Variables,

  Expectation, $E(X) = \sum_{\text{all } i} x_i \, p(x_i)$

* For contineous random variables,

  Expectation, $E(X) = \int_{-\infty}^{\infty} x \, f(x) \, dx$

* For Discrete and Contineous random variable,

  Variance, $V(X) = E\left[(X - E(X))^2\right] = \sigma^2$

  $V(X) = E(X^2) - [E(X)]^2 = \sigma^2$

* Standard deviation,

  $\sigma = \sqrt{V(X)} = \sqrt{\sigma^2}$

Note

→ E(X) is also called mean, μ, or first moment of X.

→ The quantity $E(X^n)$, $n \geq 1$ is called n^{th} moment of X, and is computed as:-

$E(X^n) = \sum_{\text{all } i} x_i^n \, p(x_i)$ → if x is discrete

$E(X^n) = \int_{-\infty}^{\infty} x^n f(x) \, dx$ → if x is contineous.

problem 3:

The mean of life of previous inspection device is $E(x) = 1/2 \int_0^\infty x\, e^{-x/2}\, dx$. Compute the variance and standard deviation.

Solution,

Mean, $E(x) = \frac{1}{2} \int_0^\infty x\, e^{-x/2}\, dx$

$$= \frac{1}{2} \left. \frac{x\, e^{-x/2}}{-1/2} \right|_0^\infty + \left. \frac{x^2}{2} e^{-x/2} \right|_0^\infty$$

$$E(x) = 2$$

Variance,

$\cancel{E(x^2) =}$

$$V(x) = E(x^2) - E(x)^2$$

$$= \frac{1}{2} \int_0^\infty x^2 e^{-x/2}\, dx - 4$$

$$= \frac{1}{2} \left. \frac{x^2 e^{-x/2}}{-1/2} \right|_0^\infty + \left. e^{-x/2} \frac{x^3}{3} \right|_0^\infty - 4$$

$$= 8 - 4$$

$$\boxed{V(x) = 4}$$

Standard Deviation,

$$\sigma = \sqrt{V(x)}$$

$$= \sqrt{4}$$

$$\boxed{\sigma = 2}$$

The Mode

* In discrete case, the mode is the value of x that occurs most frequently.
* In contineous case, the mode is the value at which the p.d.f is maximized.

DISCRETE DISTRIBUTIONS

* Used to describe random phenomenon in which only integer values can occur.
* Here, we will see :-
    1. Bernaulli Trials and Bernaulli Distributions
    2. Binomial Distribution
    3. Geometric and Negative Binomial Distributions.
    4. Poissons Distribution.

Bernaulli Trials and Bernaulli Distribution

* A Bernaulli trial is an experiment with result of success or failure.
* We can use a random variable to model this phenomenon.

  Let $X_j = 1$ if its a success

  $X_j = 0$ if its a failure.

* A consequetive 'n' Bernoulli trials are called a Bernoulli process if :-
    → Trials are independent of each other
    → Each trial has only two possible outcomes. (success / failure, True / False etc)
    → probability of a success remains constant.

$$\therefore p(x_1, x_2 \ldots x_n) = p_1(x_1) \cdot p_2(x_2) \ldots\ldots p_n(x_n)$$

$$P_j(x_j) = P(x_j) = \begin{cases} p & x_j = 1 \quad j = 1, 2 \ldots n \\ 1 - p = q & x_j = 0 \quad j = 1, 2 \ldots n \\ 0 & \text{otherwise} \end{cases}$$

* **Mean**,

$$E(X_j) = 0 \cdot q + p \cdot 1 = \boxed{p}$$

* **Variance**,

$$V(X_j) = (0^2 \cdot q + 1^2 \cdot p) - p^2$$
$$= \boxed{p(1-p)}$$

Binomial Distribution

* The random variable X ~~that~~ denotes the number of successes in 'n' Binomial trials has a binomial distribution,

$$p(x) = \begin{cases} \binom{n}{x} p^x \cdot q^{n-x} & x = 0, 1, 2 \dots n \\ 0 & \text{otherwise} \end{cases}$$

→ $P(\overbrace{SS\dots S}^{x} \overbrace{FF\dots F}^{n-x}) = p^x \cdot q^{n-x}$, $[q = 1-p]$

→ There are,

$\binom{n}{x} = \dfrac{n!}{x!\,(n-x)!}$ outcomes having the required no. of S's & F's.

* **Mean**, $E(X) = p + p + \dots + p = \boxed{np}$

Variance,

$$V(X) = pq + pq + \dots + pq = \boxed{npq}$$

Problem 4:

A production process manufactures computer chips on the average at 2% non-conforming. Everyday, a random sample of size 50 is taken from the process. If the sample contains more than two non-conforming chips, the process will be stopped. Compute the probability that the process is stopped by the sampling ~~p~~ scheme.

Solution:

Given,
The sampling process consists of $n=50$ Bernaulli trials, each with $p = 0.02$

Let,
X: Total no. of defective chips in the sample.
X has a binomial distribution given by:

$$P(x) = \begin{cases} \binom{n}{x} p^{x} \cdot q^{n-x} , & x = 0, 1 \dots 50 \\ 0 , & \text{otherwise} \end{cases}$$

$$= \binom{50}{x} (0.02)^{x} \cdot (0.98)^{50-x}$$

We have to find $P(X>2)$

$$P(X>2) = 1 - P(X \leq 2)$$

$$= 1 - \sum_{x=0}^{2} \binom{50}{x} (0.02)^{x} (0.98)^{50-x}.$$

$$= 1 - 0.92$$

$$\boxed{P(X>2) = 0.08}$$

, Mean no. of non conforming chips $\Big\}$ $E(X) = np$
$= 50(0.02)$
$= \boxed{1}$

Variance, npq
$= \boxed{0.98}$

Geometric and Negative Binomial Distributions.

* Geometric Distribution

→ Let x = no. of trials to achieve the first success in a sequence of Bernoulli trials.

x has the geometric distribution given by :-

$$p(x) = \begin{cases} q^{x-1} \cdot p & x = 1, 2 \ldots \\ 0 & \text{otherwise} . \end{cases}$$

Justification,

$x-1$ failures, followed by 1 success

$$p(\underbrace{FFF \ldots F}_{x-1} S) = q^{x-1} \cdot p .$$

→ Mean, $\boxed{E(x) = 1/p}$ → Variance: $\boxed{V(x) = \frac{q}{p^2}}$

* Negative Binomial Distribution:

→ It is the distribution of no. of trials until the k^{th} success for $k = 1, 2, \ldots$

→ If Y has a negative Binomial Distribution with parameters p and k, its distribution is

$$P(y) = \begin{cases} \binom{y-1}{k-1} \cdot q^{y-k} \cdot p^k , & y = k, k+1, \ldots \\ 0 , & \text{otherwise} . \end{cases}$$

→ Mean

$\boxed{E(Y) = k/p}$

→ Variance,

$\boxed{V(x) = kq/p^2}$

Problem 5:

40% of the assembled ink jet printers are rejected at the inspection station. Find the probability that the first ~~rejection~~ acceptable inkjet printer is the third one inspected. Also find the probability that the third inspected printer is the second acceptable printer.

Solution:

Consider each inspection as a Bernoulli trial with $q = 0.4$ and $p = 0.6$.

$$p(3) = (0.4)^2 \cdot (0.6) = \boxed{0.096}$$

$$(\because P = q^{x-1} . p)$$

To determine the probability that third printer inspected is the second acceptable printer, we use the negative Binomial Distribution -

$$p(3) = \binom{3-1}{2-1} (0.4)^{3-2} (0.6)^2 = \boxed{0.288}$$

Poisson Distribution

* It describes many random process quite well and is mathematically quiet simple.

* pmf,

$$p(x) = \begin{cases} \dfrac{e^{-\alpha} \cdot \alpha^x}{x!} & x = 0, 1 \ldots \\ 0 & \text{otherwise} \\ \text{Here}, \alpha > 1 & \end{cases}$$

* cdf,

$$F(x) = \sum_{i=0}^{x} \frac{e^{-\alpha} . \alpha^i}{i!}$$

* $\boxed{E(x) = V(x) = \alpha}$

problem 6:

A computer repair person is "beeped" each time there is a call for service. The no. of beeps per hour is known to occur in accordance with poisson distribution with mean of $\alpha = 2$ per hour. Find the probability of three beeps in the next hour. Also find the probability of two or more beeps in a 1-Hour period.

Solution:

probability of three beeps in the next hour,

$$P(3) = \frac{e^{-2} \cdot 2^3}{3!} = \boxed{0.18}$$

probability of two or more beeps in a 1-Hour period,

$$P(2 \text{ or more}) = 1 - P(0) - P(1)$$
$$= 1 - F(1)$$
$$= 1 - 0.406$$
$$= \boxed{0.594}$$

CONTINUOUS DISTRIBUTIONS

* Continuous Distributions are used to describe random phenomenon in which random variables can take any value from some interval.

* Here, we will see :-

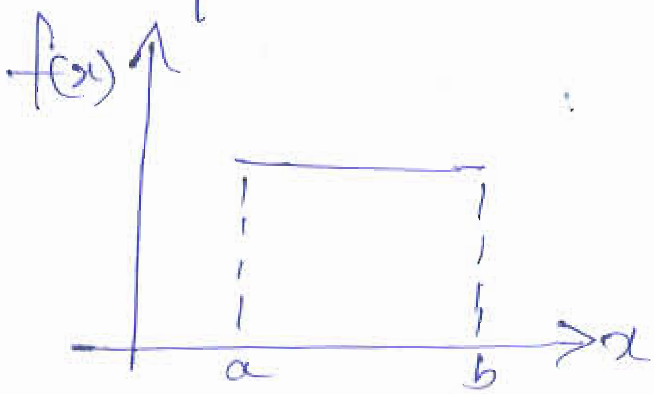
1. Uniform Distribution
2. Exponential Distribution
3. Gamma Distribution
4. Erlang Distribution
5. Normal Distribution
6. Weibull Distribution
7. Triangular Distribution
8. LogNormal Distribution
9. Beta Distribution.

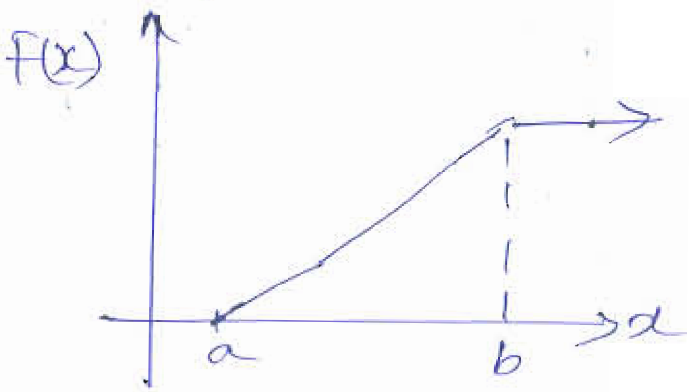
Uniform Distribution

* A random variable X is uniformly distributed on the interval (a, b) if its pdf and cdf are given by :-

$$f(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & x \geq b \end{cases}$$





NOTE

* $P(x_1 < X < x_2) = F(x_2) - F(x_1)$

$$= \frac{x_2 - x_1}{b - a}$$

* <u>Mean</u>, $E(x) = \frac{a+b}{2}$

<u>Variance</u>, $V(X) = \frac{(b-a)^2}{12}$

<u>problem 7:</u>

A Bus arrives every 20 minutes at a specified stop beginning at 6:40AM and continuing until 8:40AM. A certain passenger does not know the schedule, but arrives randomly (uniformly distributed) b/w 7:00AM and 7:30AM every morning. what is the probability that a passenger waits more than 5 mins for a bus?

<u>Solution:</u>

The passenger has to wait more than 5 min only if the arrival time is b/w 7:00AM and 7:15AM or b/w 7:20AM and 7:30AM.

Let X: Random variable that denotes no. of minutes past 7:00AM that the passenger arrives.

The desired probability is

a: 0, b= 30

$P(0 < X < 15) + P(20 < x < 30)$

$F(15) - F(0) + F(30) - F(20)$

Here $a = 0$
$b = 30$

$$\frac{15}{30} - 0 + \frac{30}{30} - \frac{20}{30} = \boxed{\frac{5}{6}}$$

Exponential Distribution

* A random variable X is said to be exponentially distributed with parameter $\lambda > 0$, if its pdf and cdf are :-

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$F(x) = \begin{cases} \int_0^x \lambda e^{-\lambda t}\,dt & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Mean,

$$E(X) = 1/\lambda$$

Variance,

$$V(X) = 1/\lambda^2$$

Note

* Exp. Dist. is used to model
  - → Completely random inter arrival times
  - → Highly variable service times
  - → Lifetime of a component that fails instantaneously.

Here, λ is ~~used~~ interpreted as a rate: ie arrivals/hour, service/minute, failure rate.

* Memoryless property of the exponential distributed random variable states that "the future values of the exponentially distributed values are not affected by the past values"

In other words, we know nothing about the future values of a random variable, given a full history of the past.

Mathematical proof:-

$$P(X > s+t \mid X > s) = \frac{P(X > s+t)}{P(X > s)}$$

$$= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}$$

$$= e^{-\lambda t}$$

$$\boxed{P(X > s+t \mid X > s) = P(X > t)}$$

problem 8:

Suppose that the life of an industrial lamp in thousands of hours, is exponentially distributed with failure rate $\lambda = 1/3$ per hour. ie on average, 1 failure per 3 hours.

a.) Find the probability that the lamp lasts longer than its mean life.

b.) Find the probability that the lamp lasts 2 to 3 hours.

c.) Find the probability that the lamp lasts another hour, given it is operating for 2.5 hours.

Solution:

a.)

$$P(X > E(X)) = P(X > 3) = 1 - P(X \leq 3)$$
$$= 1 - (1 - e^{-3/3})$$
$$= e^{-1}$$
$$= \boxed{0.368}$$

$\Rightarrow$ Independent of λ.

b.)

$$P(2 <= X <= 3) = F(3) - F(2)$$
$$= \frac{1}{e^{2/3}} - \frac{1}{e}$$
$$= \boxed{0.145}$$

c.)

$$P(X > 3.5 \mid X > 2.5) = P(X > 1)$$
$$= 1 - F(1)$$
$$= 1/e^{1/3}$$
$$= \boxed{0.717}$$

Gamma Distribution.

* A random variable X is gamma distributed with parameters β and θ if its pdf and cdf are:-

$$f(x) = \begin{cases} \dfrac{(\beta)(\theta)}{\Gamma(\beta)} (\beta\theta x)^{\beta-1} \cdot e^{-\beta\theta x} & , x>0 \\ 0 & , \text{otherwise} \end{cases}$$

$$F(x) = \begin{cases} 1 - \displaystyle\int_{\cancel{x}}^{\infty} \dfrac{\beta\theta}{\Gamma(\beta)} (\beta\theta \cdot t)^{\beta-1} \cdot e^{-\beta\theta t}\, dt & , x>0 \\ 0 & , \text{otherwise} \end{cases}$$

* $\Gamma(\beta) \rightarrow$ Gamma function

$$\Gamma(\beta) = \int_0^{\infty} x^{\beta-1} \cdot e^{-x}\, dx \quad , \quad \beta>0.$$

Crucial property,

$$\Gamma(\beta) = (\beta-1)\,\Gamma(\beta-1)$$

$\beta \rightarrow$ shape parameter

$\theta \rightarrow$ scale parameter

Note! when β is an integer, gamma distribution is related to exponential dist.

$$f(x;) = \begin{cases} \beta\theta \, e^{-\beta\theta x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Erlang Distribution

→ when $\beta = k$, an integer, we refer to the gamma distribution as the Erlang Distribution

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_k)$$

$$= k \frac{1}{k\theta} = \boxed{1/\theta}.$$

$$V(X) = V(X_1) + V(X_2) \cdots V(X_k)$$

$$= k \, 1/(k\theta)^2 = \boxed{1/k\theta^2}$$

Normal Distribution

* A random variable x has the normal distribution with mean μ and variance σ^2, if its given by the pdf :-

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty$$

* <u>Notation</u>, $X \sim N(\mu, \sigma^2)$
Random variable x is normally distributed with mean μ, and varianca σ^2.

* <u>Special properties of Normal Distribution</u>

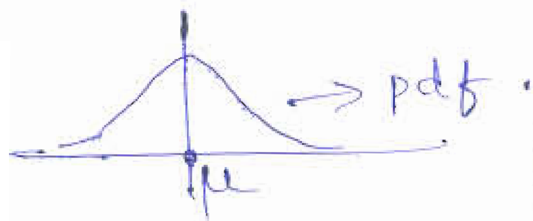
1. $\lim_{x\to-\infty} f(x) = 0$, $\lim_{x\to\infty} f(x) = 0$ — The values of $f(x)$ approaches to 0 as x tends to be infinity.

2. $f(\mu - x) = f(\mu + x) \rightarrow$ pdf is symmetric about μ.

3. Max value of pdf occurs at $x=\mu$, the mean (or mode).

* CDF,

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2} dt$$



How to evaluate this CDF ?

* Transformation of variables $z = \frac{t-\mu}{\sigma}$ allows the evaluation to be independent of μ and σ.

If $X \sim N(\mu, \sigma^2)$

Let $z = \frac{x-\mu}{\sigma}$

we get,

$$F(x) = P(X \le x) = P\left(z \le \frac{x-\mu}{\sigma}\right)$$

$$= \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

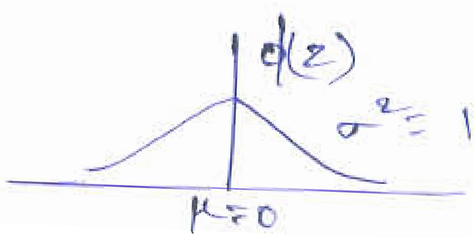
$$= \int_{-\infty}^{\frac{x-\mu}{\sigma}} \phi(z)\, dz$$

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

* the pdf, $\phi(z) = 1/\sqrt{2\pi} \cdot e^{-z^2/2} \quad -\infty < z < \infty$

is the pdf of a normal distribution with mean zero and variance 1.

* Thus, $Z \sim N(0,1)$ and it is said that z has a standard Normal distribution.



The CDF for standard normal distribution is given by :-

$$\Phi(z) = \int_{-\infty}^{z} 1/\sqrt{2\pi}\; e^{-t^2/2}\, dt.$$

Problem 9:

Given $X \sim N(50, 9)$. Compute $F(56)$

Solution:

$$F(x) = P(X \le x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$F(56) = P(X \le 56) = \Phi\left(\frac{56-50}{3}\right)$$

$$= \Phi(2)$$

$$= \int_{-\infty}^{2} 1/2\sqrt{\pi} \cdot e^{-t^2/2}\, dt$$

$$= \frac{1}{\sqrt{2\pi}}\left(\frac{e^{-t^3/2}}{-t^3/6}\right)_{-\infty}^{2}$$

$$= \sqrt{\frac{2}{\pi}}\left(-e^{-8/2} + e^{\infty}\right).$$

$$= \sqrt{0.63662} - 0.0185185$$

$$0.8 - 0.08$$

$$= \boxed{0.72}$$

Problem: (on Normal Dist)

Time to pass thou' a queue to begin self service at a cafeteria has been found to be $N(15, 9)$. Find the probability that an arriving customer waits b/w 14 and 17 min.

Solution:

$$P(14 \le x \le 17) = F(17) - F(14)$$

$$= \Phi\left(\frac{17-15}{3}\right) - \Phi\left(\frac{14-15}{3}\right)$$

$$= \Phi(0.667) - \Phi(-0.333)$$

$$= \boxed{0.3780.}$$

Weibull Distribution

* A random variable X has a weibull distribution if its pdf has the form:

$$f(x) = \begin{cases} \frac{\beta}{\alpha}\left(\frac{x-\nu}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{x-\nu}{\alpha}\right)^{\beta}\right], & x \geq \nu \\ 0, & \text{otherwise} \end{cases}$$

$\nu \to$ location parameter
$\beta \to$ scale parameter
$\alpha \to$ shape parameter

when $\nu = 0$, & $\beta = 1$:

$$f(x) = \begin{cases} \frac{1}{\alpha} e^{-x/\alpha} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Triangular Distribution

* A random variable X has a triangular distribution if its pdf is given by:

$$f(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & a \leq x \leq b \\ \frac{2(c-x)}{(c-b)(c-b)} & b \leq x \leq c \\ 0 & \text{elsewhere} \end{cases}$$

where $a \leq b \leq c$.

$$E(x) = \frac{a+b+c}{3}$$

mode, $b = 3\,E(x) - (a+c)$

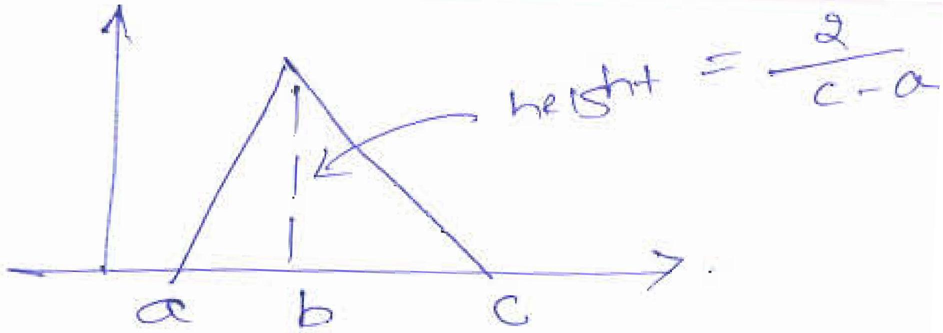


Fig: pdf of triangular Dist.

Lognormal Distribution

* A random variable x has a lognormal distribution if its pdf is given by:

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\,\sigma x} \exp\left[\dfrac{(\ln x - \mu)^2}{2\sigma^2} \right], & x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

where $\sigma^2 > 0$.

$$E(x) = e^{\mu + \sigma^2/2}$$

$$V(x) = e^{2\mu + \sigma^2} \cdot \left(e^{\sigma^2} - 1\right)$$

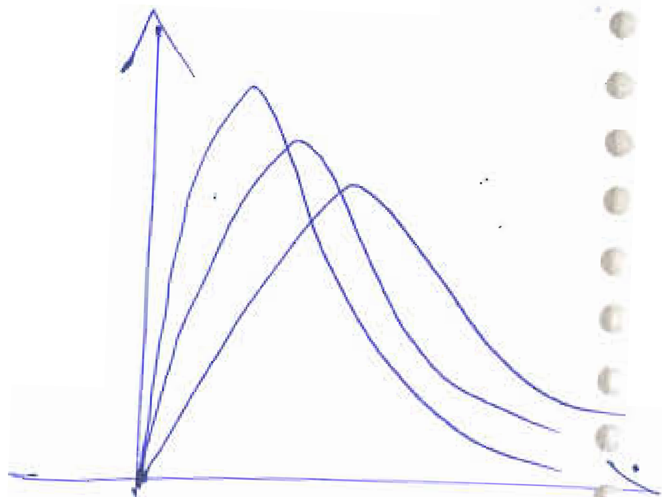


Fig: PDF of LogNormal Distribution.

POISSON PROCESS

* <u>Definition:</u>

$N(t)$ is a counting function that represents the number of events occurred in $(0,t)$

A counting process $\{N(t), t \geq 0\}$ is a <u>poisson process</u> with mean rate λ if

1. Arrivals occur one at a time.
2. $\{N(t), t \geq 0\}$ has stationary increments.
3. $\{N(t), t \geq 0\}$ has independent increments

→ Distribution of no. of ~~intervals~~ arrivals b/w t and $t+s$ depends only on length of interval s, and not on starting point t. Thus arrivals are completely at random.

→ Number of arrivals during non overlapping time intervals are independent random variables. Thus a large or small number of arrivals in one time interval has no effect on the number of arrivals in subsequent time intervals.

If arrivals occurs according to a poissons process, it can be shown that $N(t)$ is equal to n is given by:-

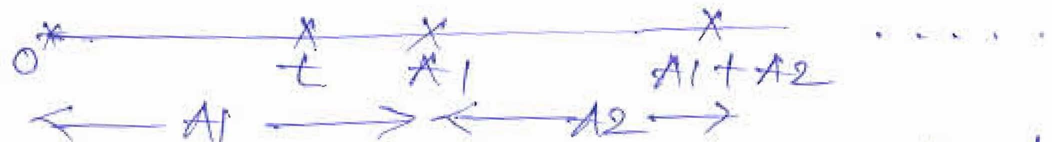
$$P[N(t)=n] = \frac{e^{-\lambda t} \cdot (\lambda t)^n}{n!}$$

$t \geq 0$
$n = 0, 1, 2 \ldots\ldots$

Mean,
$E(x) = \lambda t$
Variance
$V(x) = \lambda t$

Inter Arrival times

* Consider the inter arrival times of a poisson process, $(A_1, A_2 \ldots\ldots)$
  $A_i \rightarrow$ elapsed time b/w arrival i and $i+1$



$\Rightarrow$ First Arrival occurs after time t if and only if there are no arrivals in $(0, t)$. Hence -

$P(A_1 > t) = P(N(t) = 0) = e^{-\lambda t}$

$P(A_1 \leq t) = \underline{\underline{1 - e^{-\lambda t}}} \rightarrow$ probability that the first arrival will occur in $(0, t)$.

$\Rightarrow$ Inter arrival times $A_1, A_2 \ldots$ are exponentially distributed and independent with mean $1/\lambda$.

Properties of a poissons process

1. Splitting
2. Pooling.

Splitting

* Suppose each event of a poisson process can be classified as type I with probability p and type II with probability $1-p$

$N(t) = N_1(t) + N_2(t)$
where $N_1(t)$ and $N_2(t)$, both are poissons processes with rates λp and $\lambda(1-p)$.

$$N(t) \sim poi(\lambda) \xrightarrow{\lambda} \begin{cases} \xrightarrow{\lambda p} N_1(t) \sim poi(\lambda p) \\ \xrightarrow{\lambda(1-p)} N_2(t) \sim poi(\lambda(1-p)) \end{cases}$$

Pooling

* Two poissons processes can be pooled together. $N_1(t) + N_2(t) = N(t)$ where. $N(t)$ is a poissons process with rates $\lambda_1 + \lambda_2$

$$N_1(t) \sim poi(\lambda_1) \xrightarrow{\lambda_1}$$
$$\xrightarrow{\lambda_1+\lambda_2} N(t) \sim poi(\lambda_1 + \lambda_2)$$
$$N_2(t) \sim poi(\lambda_2) \xrightarrow{\lambda_2}$$

Non stationary poissons process (NSPP)

* NSPP is a poissons process without the stationary increments, characterized by $\lambda(t)$, arrival rate at time t.
* The expected number of arrivals by time t, $\Lambda(t)$;

$$\boxed{\Lambda(t) = \int_0^t \lambda(s)\,ds}$$

problem:

Suppose ~~that~~ Arrivals to a post office occur at a rate of 2 per minute from 8AM until 12 NOON and then drops to 1 every 2 minutes until the day ends at 4PM. what is the probability ~~that~~ distribution of no. of arrivals b/w 11AM and 2PM?

Solution:

Let $t=0$ corresponds to 8AM.

NSPP $N(t)$ has rate function,

$$\lambda(t) = \begin{cases} 2 & 0 \le t < 4 \\ 1/2 & 4 \le t \le 8 \end{cases}$$

Expected number of arrivals by time t is therefore,

$$\Lambda(t) = \begin{cases} 2t & 0 \le t \le 4 \\ t/2 + 6 & 4 \le t \le 8 \end{cases}$$

↳ we got this by applying $\Lambda(t) = \int_0^t \lambda(s)\,ds$

Hence the prob. distribution of the no. of arrivals b/w 11AM and 2PM is

$$P(N(6) - N(3) = k) = P(N(\Lambda(6)) - N(\Lambda(3)) = k)$$

$$= P[N(9) - N(6)]$$

$$= \frac{e^{9-6} \cdot (9-6)^k}{k!}$$

$$= \boxed{\frac{e^{3} \cdot 3^k}{k!}}$$

Ashok Kumar . K

9742024066

celestialcluster@gmail.com

UNIT 5

RANDOM-NUMBER GENERATION, RANDOM-VARIATE GENERATION

Syllabus

RANDOM NUMBER GENERATION

* properties of Random Numbers
* Generation of pseudo-Random numbers
* Techniques for generating Random numbers
* Tests for Random numbers.

RANDOM VARIATE GENERATION

* Inverse transform technique
* Acceptance-Rejection technique
* Special properties.

– 8 Hours.

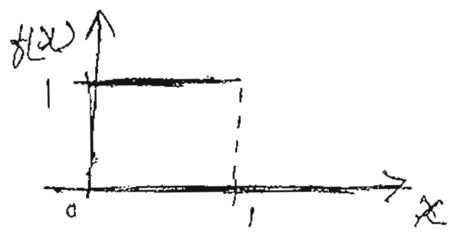
* Random numbers are a necessary basic ingredient in the simulation of almost all discrete systems.
Most computer languages have a subroutine, object, or function that will generate a random number.
Similarly, simulation languages generate random numbers ~~that~~ are used to generate event times and other random variables.

PROPERTIES OF RANDOM NUMBERS

* A sequence of random numbers $R_1, R_2, \ldots$ must have two important statistical properties:
  - → uniformity
  - → Independence.

Each random number R_i must be an independent sample drawn from a continuous uniform distribution between zero and one.
ie the pdf is given by

$$\boxed{f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}}$$



This density function is shown in above fig:
The expected value of each R_i is given by:

$$E(R) = \int_0^1 x\,dx = \frac{x^2}{2}\Big|_0^1 = \boxed{\frac{1}{2}}.$$

and the variance is given by:

$$V(R) = \int_0^1 x^2 dx - [E(R)]^2 = \boxed{\frac{1}{12}}$$

* Some consequences of the uniformity and independence properties are ;
  - → If the interval [0, 1] is divided into 'n' classes, or subintervals of equal length, the expected number of observations in each intervals is N/n, where $N \rightarrow$ total no. of observations.
  - → The probability of observing a value in a particular interval is independent of the previous values drawn.

GENERATION OF PSEUDO RANDOM NUMBERS

↳ means "false"

* "pseudo", because generating numbers using a known method removes the potential for true randomness.
* The goal of any generation scheme is to produce a sequence of numbers b/w 0 and 1 that simulates, or imitates, the ideal properties of random no's (RN)

~~Important considerations~~

* In the generation of pseudo RN, certain problems or errors can occur. Some ex. for such errors (or departure from ideal randomness) are :
  - → The generated nos might not be uniformly distributed.
  - → They may be discrete-valued instead of continuous-valued.
  - → The mean of generated no's may be too high or too low.

→ The variance of the generated no's might be too high or too low.

→ There might be dependence, the following are ex:

(a) autocorrelation b/w nos.

(b) no's successively higher or lower than adjascent no's.

(c) several no's above the mean followed by several no's below the mean.

* <u>Important considerations in RN routines are</u>:

1.) The routine should be <u>fast</u>. Individual computations are inexpensive, but simulation could require many millions of random numbers.

2.) The routine should be <u>portable</u> to different computers — and, ideally to different <u>programming</u> languages.

3.) The routine should have a sufficiently <u>long cycle</u>. The <u>cycle length</u>, or <u>period</u>, represents the length of random number sequence before previous numbers begin to repeat themselves in an earlier order.

4) The random numbers should be <u>replicable</u>. Given the starting point it should be <u>possible</u> to generate the same set of RN, completely independent of the system that is being simulated.

*5.) Generated RN should closely approximate the ideal statistical properties of uniformity and independence.

TECHNIQUES FOR GENERATING RANDOM NUMBERS

1. Linear Congruential Method (LCM) → most widely used.
2. Combined Linear Congruential Generators (CLCG)
3. Random number streams

Linear Congruential method (LCM)

* proposed by Lehmar [1951]
* LCM produces a sequence of integers, $x_1, x_2 \ldots$ between 0 and m-1 following a recursive relnship:

$$\boxed{X_{i+1} = (aX_i + c) \bmod m} \quad i = 0, 1, 2 \ldots$$

X_0 → intial value, & is called seed
a → multiplier.
c → increment.
m → modulus.

~~* If c=0~~

* If $c \neq 0$, then the form is called mixed congruential method. When $c = 0$, the form is called multiplicative congruential method.
* selection of values of a, c, m, & x_0 affects the statistical properties & cycle length.
* Random numbers between 0 and 1 can be generated by

$$\boxed{R_i = \frac{X_i}{m}} \quad i = 1, 2 \ldots$$

* Example 1:
Use the LCM to generate a sequence of RN with $x_0 = 27$, $a = 17$, $c = 43$, and $m = 100$.

Solution:-

$X_0 = 27$.

$X_1 = (a \cdot X_0 + c) \bmod m$

$= (17 \times 27 + 43) \bmod 100$

$= 502 \bmod 100$

$\boxed{X_1 = 2} \Rightarrow R_1 = \frac{2}{100} = \boxed{0.02}$.

$X_2 = (2 \times 17 + 43) \bmod 100$.

$\boxed{X_2 = 77} \Rightarrow \boxed{R_2 = 0.77}$.

$X_3 = (17 \times 77 + 43) \bmod 100$

$\boxed{X_3 = 52} \Rightarrow \boxed{R_3 = 0.52}$.

etc

* Other than uniformity and independence, there are several secondary properties that must be considered.
these include. maximum density & maximum period.

→ By maximum Density is meant that the values assumed by R_i, $i = 1, 2 \ldots$ leave no large gaps on $[0, 1]$

→ To help achieve max. density, and to avoid cycling in practical applications, the generator should have the largest possible period.

* Maximal period can be achieved by ~~the~~ proper choice of a, c, m, and x_0.

(i) For m a power of 2, say $m = 2^b$ and $c \neq 0$, the longest possible period is $P = m = 2^b$, which is achieved whenever 'c' is relatively prime to 'm' (ie ~~gcd~~ ~~greatest common factor~~ of m & c is 1) and $a = 1 + 4k$. where $k \rightarrow$ integer.

(ii) For m a power of 2, say $m = 2^b$ and $c = 0$. the longest possible period is $P = m/4 = 2^{b-2}$, which is achieved if the seed x_0 is odd and if the multiplier, 'a', is given by $a = 3 + 8k$ or $a = 5 + 8k$.

(iii) For m a prime number and $c = 0$, the longest possible period is $P = m - 1$, which is achieved whenever the multiplier, 'a', has the property that the smallest integer 'k' such that $a^k - 1$ is divisible by m is $k = m - 1$.

* Example 2:

Using the Multiplicative congruential method, find the period of the generator for $a = 13$, $m = 2^6 = 64$ and $X_0 = 1, 2, 3, \& 4$.

Solution.

i	X_i	X_i	X_i	X_i
0	1	2	3	4
1	13	26	39	52
2	41	18	59	36
3	21	42	63	20
4	17	34	51	4
5	29	58	23	
6	57	50	43	
7	37	10	47	
8	33	2	35	
9	45		7	
10	9		27	
11	53		31	
12	49		19	
13	61		55	
14	25		11	
15	5		15	
16	1		3	

Here $m = 2^6 = 64$ and $c = 0$.

Therefore,

The maximal period is $P = \frac{m}{4} = \boxed{16}$.

This ~~is~~ period is achieved by using odd seeds $X_0 = 1$ & $X_0 = 3$.

even seeds $X_0 = 2$ and $X_0 = 4$ yield the periods 8 and 4 respectively, both less than the maximum.

Example 3:

let $m = 10^2 = 100$, $a = 19$, $c = 0$, $X_0 = 63$.

Generate a sequence of Random integers.

soln:

$X_0 = 63$

$X_1 = (19)(63) \bmod 100 = 97$

$X_2 = (19)(97) \bmod 100 = 43$

$X_3 = (19)(43) \bmod 100 = 17$

⋮

Combined Linear Congruential Generators

* Reason: longer period generation is needed because of increasing complexity of simulated systems.
  Approach: Combine two or more multiplicative congruential generators.

* If $W_{i,1}, W_{i,2} \dots W_{i,k}$ are any independent, discrete valued random variables, and one of them, say $W_{i,1}$, is uniformly distributed on the integers from 0 to $m_1 - 2$, then

$$W_i = \left[\sum_{j=1}^{k} W_{i,j}\right] \bmod (m_1 - 1)$$ is uniformly

distributed on the integers from 0 to $m_1 - 2$.

* using this result, we can form a combined generators;
  let $X_{i,1}, X_{i,2} \dots X_{i,k}$ be the i^{th} output from k different multiplicative congruential generators, where the j^{th} generator ~~has prime modulus m_j and the multiplier a_j~~
  → has prime modulus m_j & multiplier a_j & period $m_j - 1$
  → produces integers $X_{i,j}$ that are approximately uniformly distributed on integers from 1 to $m_j - 1$
  → $W_{i,j} = X_{i,j} - 1$ is approximately uniformly distributed on the integers from 0 to $m_j - 2$

* Combined generators are of the form:

$$X_i = \left[\sum_{j=1}^{k} (-1)^{j-1} X_{i,j} \right] \bmod m_1 - 1$$

with

$$R_i = \begin{cases} X_i/m_1 , & X_i > 0 \\ \dfrac{m_1 - 1}{m_1} , & X_i = 0 \end{cases}$$

$$\left.\begin{array}{r}\text{maximum}\\ \text{possible}\\ \text{period}\end{array}\right\} P = \frac{(m_1-1)(m_2-1)\cdots(m_k-1)}{2^{k-1}}$$

* Example ~~problem~~: For a 32 bit computers, L'Ecuyer suggests combining $k=2$ generators with $m_1 = 2{,}147{,}483{,}563$, $a_1 = 40{,}014$, $m_2 = 2{,}147{,}483{,}399$, and $a_2 = 40{,}692$. This leads to following algorithm

1. Select seed $X_{1,0}$ in the range $[1, 2147483562]$ for the first generator, and seed $X_{2,0}$ in the range $[1, 2147483398]$ for the second

   Set $j = 0$

2. Evaluate each individual generator.

$$X_{1,j+1} = 40{,}014\, X_{1,j} \bmod 2147483563$$

$$X_{2,j+1} = 40692\, X_{2,j} \bmod 2147483399$$

3. Set

$$X_{j+1} = (X_{1,j+1} - X_{2,j+1}) \bmod 2{,}147{,}483{,}562$$

4. Return

$$R_{j+1} = \begin{cases} \dfrac{X_{j+1}}{2147483563}, & X_{j+1} > 0 \\ \dfrac{2147483562}{2147483563}, & X_{j+1} = 0. \end{cases}$$

5. set $j = j+1$ and go to step 2.

This combined generator has period

$$\frac{(m_1 - 1)(m_2 - 1)}{2} \approx \underline{2 \times 10^{18}}$$

Random Number Streams.

TESTS FOR RANDOM NUMBERS

* The tests can be placed in two categories, according to the properties of interest
  - → uniformity
  - → Independence.

1.) Testing for uniformity.

Here, the hypotheses are as follows:

$H_0 : R_i \sim U[0,1]$

$H_1 : R_i \nsim U[0,1]$

The null Hypothesis H_0, reads that the numbers are distributed uniformily on the interval $[0,1]$
Failure to reject the null hypothesis means that evidence of nonuniformity has not been detected by this test.
This does not imply that further testing of the generator for uniformity is unnecessary.

2.) Testing for independence.

Here, the hypothesis are as follows:

$H_0 : R_i \sim$ independently

$H_1 : R_i \nsim$ independently

This null Hypothesis H_0, reads that the numbers are independent failure to reject the null hypothesis means that evidence of dependence has not been detected by this test. This does not imply that further testing of generator for independence is unnecessary.

* For each test, a level of significance α must be stated. The level α is the probability of rejecting the null hypothesis when the null hypothesis is true:

$$\boxed{\alpha = P(\text{reject } H_0 \mid H_0 \text{ true})}$$

* When to use these tests:
  - → If a well known simulation languages or random-numbers generators are used, it is probably unnecessary to test.
  - → If the generator is not explicitly known or documented, eg: spreadsheet programs, tests should be applied to many sample numbers.

* Types of tests:
  - → Theoritical tests: Evaluate the choices of m, a, & c without actually generating any numbers.
  - → Empirical tests: Applied to actual sequence of numbers produced.

1. Frequency tests.

* A basic test that should always be performed to validate a new generator is the test of uniformity. Two different methods of testing are available
  - → Kolmogorov - Smirnov test
  - → chi-square test.

The Kolmogorov-Smirnov test.

* This test compares the continuous cdf, $F(x)$ of the uniform distribution with the empirical cdf, $S_N(x)$, of the sample of N observations.

By Definition,

$$F(x) = x \quad , \quad 0 \le x \le 1$$

If the sample from the random-number generator is $R_1, R_2 \ldots R_N$, then the empirical cdf $S_N(x)$ is defined by

$$S_N(x) = \frac{\text{number of } R_1, R_2 \ldots R_N \text{ which are } \le x}{N}.$$

As N becomes larger, $S_N(x)$ should become a better approximation to $F(x)$, provided that the null hypotheses is true.

* This test is based on the largest absolute deviation b/w $F(x)$ and $S_N(x)$ over the range of the random variable - i.e, it is based on the statistic:

$$D = \max \left| F(x) - S_N(x) \right|$$

* For testing against a uniform cdf, the test procedure follows these steps.

Step 1: Rank the data from smallest to largest. let $R_{(i)}$ denote the i^{th} smallest observation, so that

$$R_{(1)} \le R_{(2)} \le \cdots \le R_{(N)}$$

Step 2: Compute

$$D^+ = \max_{1 \le i \le N} \left\{ \frac{i}{N} - R_{(i)} \right\}$$

$$D^- = \max_{1 \le i \le N} \left\{ R_{(i)} - \frac{i-1}{N} \right\}$$

Step 3: Compute $D = \max(D^+, D^-)$

Step 4: locate in table (K-S critical values) the critical value D_α for the specified significance level α and the given sample size N. given in next page,

Step 5: If the sample statistic D is greater than the critical value D_α, the null hypothesis that the data are a sample from a uniform distribution is rejected.

if $D \le D_\alpha$, conclude that no difference has been detected b/w the true distribution of $\{R_1, R_2 \ldots R_N\}$ and the uniform distribution.

Example:

Suppose the 5 generated nos are 0.44, 0.81, 0.14, 0.05, 0.93.

step 1:

$R_{(i)}$	0.05	0.14	0.44	0.81	0.93
i/N	0.20	0.40	0.60	0.80	1.00
$i/N - R_{(i)}$	0.15	0.26	0.16	—	0.07
$R_{(i)} - (i-1)/N$	0.05	—	0.04	0.21	0.13

step2:

$D^+ = \max\{i/N - R_{(i)}\}$

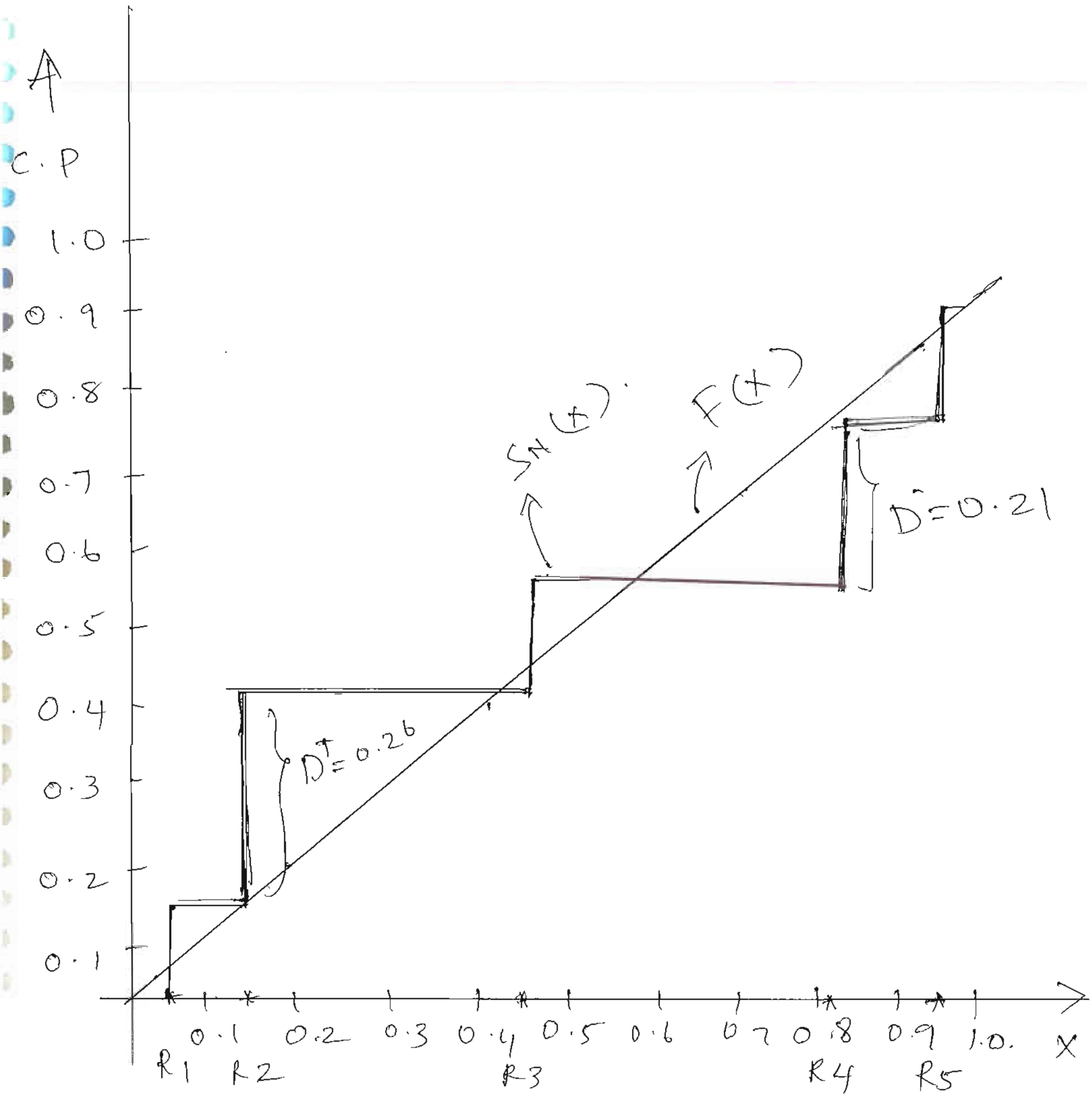
$D^- = \max\{R_i - (i-1)/N\}$

step 3: $D = \max(D^+, D^-) = 0.26$

step 4: For $\alpha = 0.05$,
$D_\alpha = 0.565 > D$

~~step~~ Hence, H_0 is not rejected.

comparing $F(x)$ and $S_N(x)$ on a graph.



~~The chi square Test~~

ref (Fig.) in text book

The chi-square test

* It uses the sample statistic

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$O_i \rightarrow$ observed no. in the i^{th} class.

$E_i \Rightarrow$ Expected number in the i^{th} class

$n \rightarrow$ no. of classes.

For ~~unif~~ the uniform distribution, E_i is given by

$$E_i = \frac{N}{n} \text{ for equally spaced classes.}$$

$N \rightarrow$ total no. of observations.

for ex problem
after text book.

Problem:

use the chi-square test with $\alpha = 0.05$ to test for whether the data shown are, uniformly distributed in ten intervals $(0, 0.1)$, $(0.1, 0.2)$ ..... $(0.9, 1)$

Material	O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
1	8	10	-2	4	0.4
2	8	10	-2	4	0.4
3	10	10	0	0	0
4	9	10	-1	1	0.1
5	12	10	2	4	0.4
6	8	10	-2	4	0.4
7	10	10	0	0	0
8	14	10	4	16	1.6
9	10	10	0	0	0
10	11	10	1	1	0.1
					3.4

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(E_i - O_i)^2}{E_i}$$

$$= 3.4$$

$$\chi_0^2 \leq \chi^2_{\alpha, n-1}$$

$3.4 \leq 16.9$. TRUE

$\therefore$ H_0 is NOT REJECTED.

Tests for Auto correlation

* Tests for auto correlation are concerned with the dependence b/w nos: in a sequence.
* This test requires the computation of autocorrelation b/w every m nos (m is a.k.a the lag), starting with the i^{th} no.
  Thus, the auto correlation ρ_{im} b/w the following nos would be of interest: $R_i; R_{i+m}, R_{i+2m} \dots R_{i+(M+1)m}$
  The value M is the largest integer such that $i+(M+1)m \le N$, where N is the total no. of values in the sequence.
* A nonzero autocorrelation implies a lack of independence so the following two-tailed test is appropriate.

  $H_0: \rho_{im} = 0$, if nos are independent
  $H_1: \rho_{im} \neq 0$, if nos are dependent.

* For large values of M, the distribution of the estimator of ρ_{im} denoted by $\hat{\rho}_{im}$ is approximately normal.
* Test statistics can be formed as:

$$Z_0 = \frac{\hat{\rho}_{im}}{\sigma_{\hat{\rho}_{im}}}$$

CHAPTER 5B

RANDOM VARIATE GENERATION.

Here we discuss techniques to produce random variables of other distributions

All techniques here assume the existence of a source of uniform (0,1) random numbers $R_1, R_2 \ldots\ldots$

pdf, $$f(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

cdf, $$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

Here, we see

1. INVERSE TRANSFORM TECHNIQUE
2. ACCEPTANCE REJECTION TECHNIQUE

INVERSE TRANSFORM TECHNIQUE

* Inverse Transform technique can be used to sample from <u>exponential</u>, <u>uniform</u>, <u>weibull</u> and the <u>triangular distributions</u>.

* The basic principal is to find the inverse function of F, F^{-1} such that

$$F \cdot F^{-1} = F^{-1} \cdot F = 1$$

R is generated from uniform (0,1)

* F^{-1} denotes the solution of the equation $R = F(X)$ in terms of R. (not $1/F$)

Eg: → Inverse of $y = x$ is $x = y$.

→ Inverse of $y = 2x + 1$ is $x = \frac{y-1}{2}$

Exponential Distribution

* pdf,

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

cdf,

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

* The idea is to solve $y = 1 - e^{-\lambda x}$ for x where y is uniformly distributed on $(0, 1)$.

- Steps involved are as follows.

Step 1:

Compute the cdf of the desired random variable X.

For exponential distribution, cdf is

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

Step 2:

Set $R = F(x)$ on the range of x.

For exponential distribution,

$\boxed{R = 1 - e^{-\lambda x}}$ on the range of $x \geq 0$

Step 3:

Solve the equation $F(x) = R$ for x in terms of R

ie $1 - e^{-\lambda x} = R$

$$e^{-\lambda x} = 1 - R$$

$$\boxed{x = \frac{-1}{\lambda} \ln(1-R)}$$

$$\boxed{x = F^{-1}(R)}$$

Step 4:

Generate the uniform random numbers (as needed) $R_1, R_2 \ldots$ and compute the Random variates by:

$$X_i = F^{-1}(R_i)$$

$$X_i = -1/\lambda \ln(1 - R_i)$$

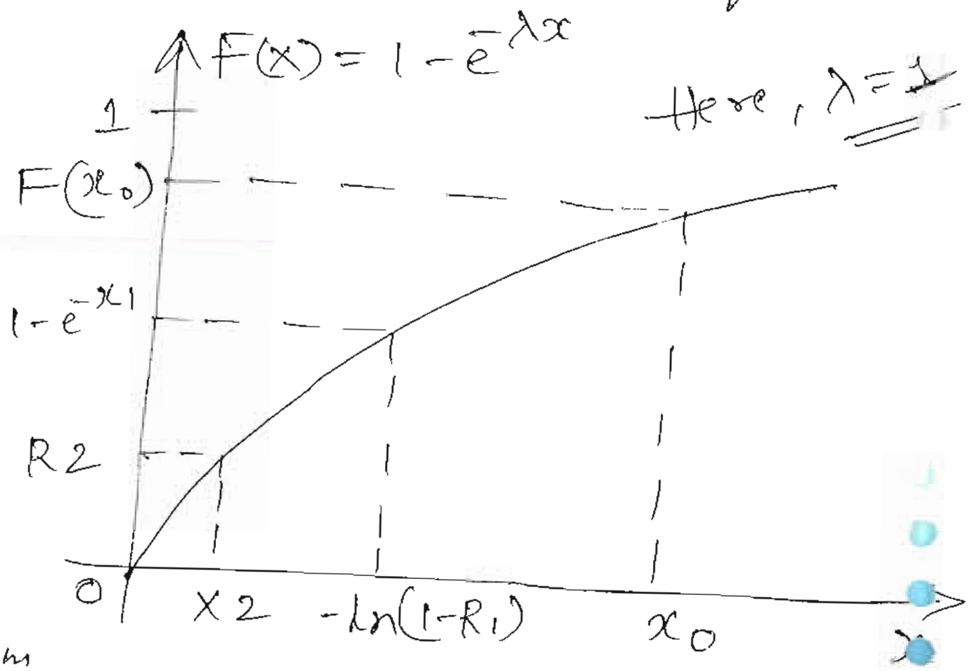
$\boxed{X_i = -1/\lambda \ln R_i}$ $i = 1, 2, 3 \ldots$

($\because$ both R_i and $1-R_i$ are uniformly distributed)

Graphical Representation of the inverse Transform technique.

- Here the cdf shown is $F(x) = 1 - e^{-x}$ an exp. dist. with $\lambda = 1$. (rate)

- To generate a value x_1 with cdf $F(x)$, a random no R_1 b/w 0 and 1 is generated, then a horizontal line is drawn from R_1 to the graph of cdf, and then a vertical line is drawn to x-axis to obtain x_1, the desired result.



Triangular Distribution

Step 1: cdf,

$$F(x) = \begin{cases} 0, & x < 0 \\ x^2/2, & 0 < x \le 1 \\ 1 - \frac{(2-x)^2}{2}, & 1 < x \le 2 \\ 1, & x > 2 \end{cases}$$

Step 2: $F(x) = R$

Step 3: Solve for x in terms of R.

$$R = \frac{x^2}{2} \Rightarrow x = \sqrt{2R} \quad \text{——— (1)}$$

$$R = 1 - \frac{(2-x)^2}{2} \Rightarrow x = 2 - \sqrt{2(1-R)} \quad \text{—— (2)}$$

Combining (1) and (2),

$$\boxed{x = \begin{cases} \sqrt{2R}, & 0 \le R \le 0.5 \\ 2 - \sqrt{2(1-R)}, & 0.5 \le R \le 1 \end{cases}}$$

Uniform Distribution

- Consider a Random variable X that is uniformly distributed on the interval (a, b). A reasonable guess for generating X is given by:

$$\boxed{X = a + (b-a)R}$$

proof:

step 1:

cdf,
$$F(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \le x \le b \\ 1 & x > b \end{cases}$$

step 2:

set
$$F(x) = \frac{x-a}{b-a} = R$$

step 3

solve for X in terms of R

$$\boxed{X = a + (b-a)R}$$

Weibull Distribution.

step 1: cdf, $F(x) = 1 - e^{-(x/\alpha)^\beta}$, $x \geq 0$.

step 2: set $F(x) = 1 - e^{-(x/\alpha)^\beta} = R$

step 3: solve for $\cancel{X}$ in terms of R,

$$\boxed{x = \alpha\left[-\ln(1-R)\right]^{1/\beta}}$$

Empirical Continuous Distribution

* Used when theoritical distributions are not applicable.
* <u>Steps involved:</u>

1. Collect Empirical data and group them accordingly.
2. Tabulate the frequency and cumulative frequency.
3. Now assume the value of cumulative freq as a function of empirical data,

$$F(x) = R$$

4. Establish a relation b/w x and R using linear Interpolation.

$$\boxed{X = F^{-1}(R) = X_{i-1} + a_i\left(R - \frac{i-1}{n}\right)}$$

$$\frac{i-1}{n} < R \le \frac{i}{n}$$

slope,

$$a_i = \frac{x_i - x_{i-1}}{i/n - (i-1)/n}$$

$$\boxed{a_i = \frac{x_i - x_{i-1}}{1/n}}$$

Problem:

Five observations of fire-crew response times (in min) to incoming alarms are given below

2.76, 1.83, 0.80, 1.45, 1.24.

Generate the random variates from the response time distribution. Given the random number 0.71.

Solution:

Step 1: Group the Empirical data in a sorted manner.

0.80, 1.24, 1.45, 1.83, 2.76.

step 2:

i	Interval $x_{i-1} < x \leq x_i$	P $1/n$	C.P i/n	slope a_i
1	$0 < x \leq 0.80$	0.20	0.20	4.00
2	$0.80 < x \leq 1.24$	0.20	0.40	2.20
3	$1.24 < x \leq 1.45$	0.20	0.60	1.05
4	$1.45 < x \leq 1.83$	0.20	0.80	1.90
5	$1.83 < x \leq 2.76$	0.20	1.00	4.65

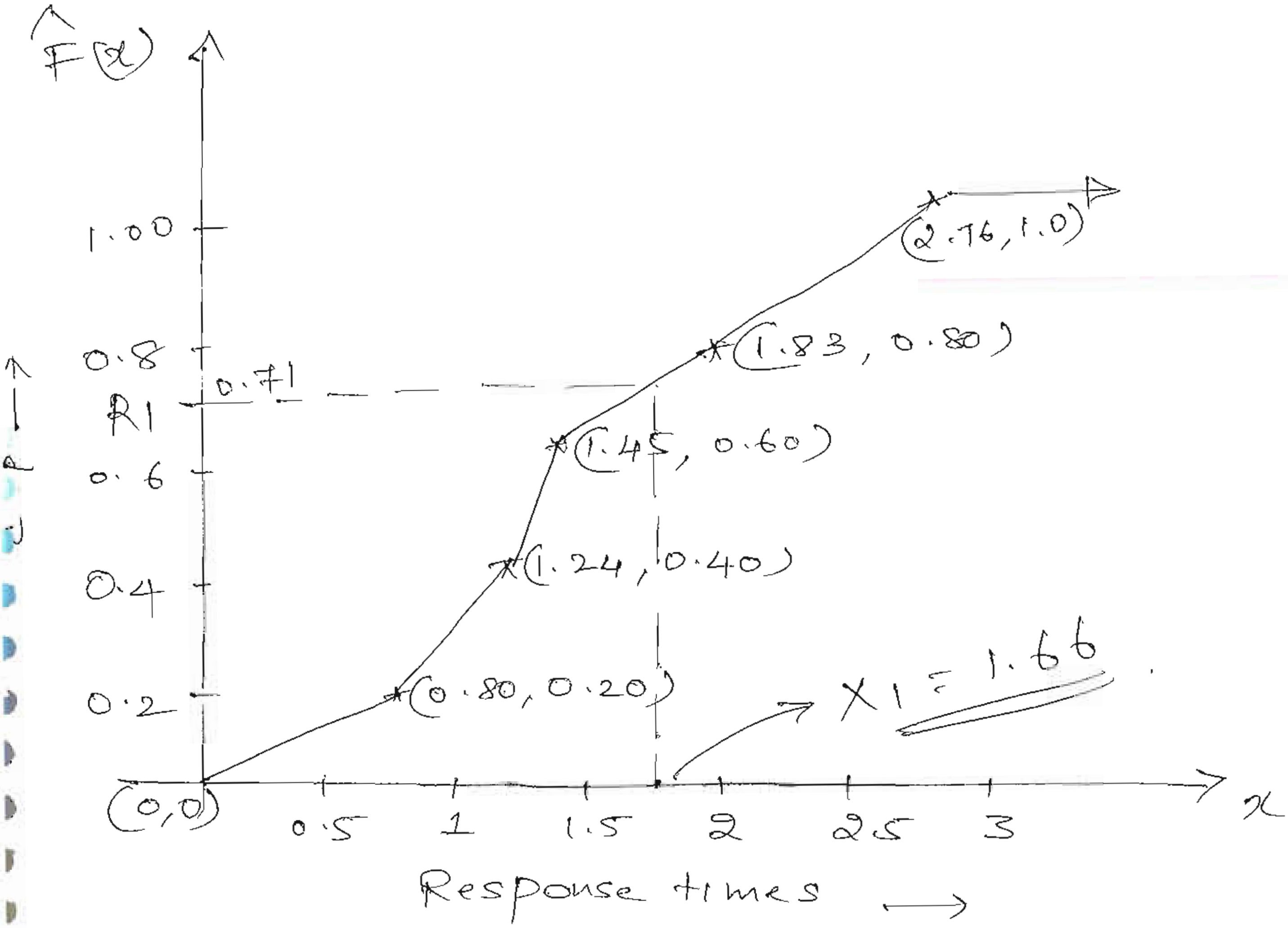
step 3:

$F(x) = R$

step 4:

Given $R_1 = 0.71$.

Below fig shows the empirical cdf of fire-crew response times.



step 4:

Given $R_1 = 0.71$ lies in the fourth interval b/w $3/5 (= 0.60)$ and $4/5 (= 0.80)$

$$\therefore \quad x_1 = x_{(4-1)} + a_4 \left[R_1 - \left(\frac{4-1}{n} \right) \right]$$

$$= 1.45 + 1.90 (0.71 - 0.60)$$

$$\boxed{x_1 = 1.66}$$

Discrete Distribution

* All discrete distributions can be generated using the inverse transform technique.

Here we discuss
- → empirical
- → uniform (discrete)
- → geometric

Example 1: An Empirical Discrete Distribution.

problem: suppose the no. of shipments X on the loading dock of the IHW company is either 0, 1, or 2 at the end of the day, with observed relative frequency of occurrence of 0.50, 0.30, & 0.20 respectively, Generate the values X, to represent the no. of shipments on the loading dock at the end of each day.

Solution:

pmf, $P(X)$

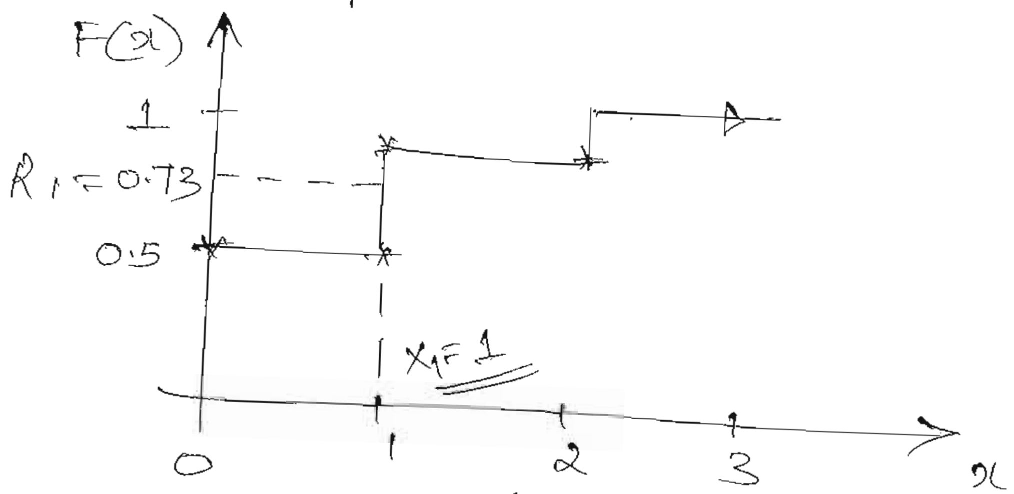
$$p(0) = P(X=0) = 0.50$$
$$p(1) = P(X=1) = 0.30$$
$$p(2) = P(X=2) = 0.20$$

Distribution of no. of shipments X

x	$P(x)$ (P)	$F(x)$ (CP)
0	0.50	0.50
1	0.30	0.80
2	0.20	1.00

cdf, $F(x) = P(X \le x)$ is given by

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.5 & 0 \le x < 1 \\ 0.8 & 1 \le x < 2 \\ 1.0 & 2 \le x \end{cases}$$



Consider $R_1 = 0.73$

$$F(x_{i-1}) < R <= F(x_i)$$

$$F(x_0) < 0.73 <= F(x_1)$$

$$\Rightarrow \boxed{x_1 = 1}$$

Since $r_1 = 0.5 < R_1 = 0.73 < r_2 = 0.8$,

Set $x_1 = x_2 = 1$

The generated scheme is summarized as,

$$x = \begin{cases} 0 & R \le 0.5 \\ 1 & 0.5 < R \le 0.8 \\ 2 & 0.8 < R \le 1.0 \end{cases}$$

i	I/P r_i	O/P x_i
1	0.50	0
2	0.80	1
3	1.00	2

← Table for generating the discrete variate x.

Example 2: A Discrete Uniform Distribution.

* pdf, $p(x) = 1/K \quad x = 1, 2 \ldots K$.

cdf,

$$F(x) = \begin{cases} 0 & x < 1 \\ i/K & i \le x < K \quad \text{for } 1 \le i \le K \\ 1 & K \le x \end{cases}$$

* Let $F(x) = R$

Solve x in terms of R.

Since x is discrete,

$$r_{i-1} \left(= \frac{i-1}{K}\right) < R \le r_i \left(= \frac{i}{K}\right)$$

$\therefore$

$$i - 1 < RK \le i$$

$$\boxed{RK \le i < RK + 1}$$

Since i and K are integers and R is b/w 0 and 1, for the above eq to hold, we need

$$\boxed{x = \lceil RK \rceil}$$

* Ex: Given the uniform distribution on $(1, 2 \ldots K)$ with pmf $p(x) = 1/K \quad x = 1, 2 \ldots K$. Generate the random variates for the five random no.s (0.81, 0.12, 0.34, 0.56, & 0.93) use $K = 10$

Derive the formula used

[June/July 08 - 10 marks]

Solution:

$R_1 = 0.81$

$\therefore \; x_1 = \lceil R K \rceil$

$= 0.81 * 10 = \lceil 8.1 \rceil \Rightarrow \boxed{R1 = 9}$

III'ly

$R_2 = 0.12 \quad \boxed{x_2 = 2}$

$R_3 = 0.34 \quad \boxed{x_3 = 4}$

$R_4 = 0.56 \quad \boxed{x_4 = 6}$

$R_5 = 0.93 \quad \boxed{x_5 = 10}$

Example 3: Geometric Distribution.

* pmf $= p(x) = p*(1-p)^x, \quad x = 0, 1, 2 \ldots$

where $0 < p < 1$

cdf, $F(x) = \sum_{j=0}^{x} p(1-p)^j$

$F(x) = 1 - (1-p)^{x+1}$

* solving for x in terms of R,

$$x = \left\lceil \frac{\ln(1-R)}{\ln(1-p)} - 1 \right\rceil$$

$$\boxed{x = \lceil -\beta \ln(1-R) - 1 \rceil}$$

where $\beta = \dfrac{-1}{\ln(1-p)}$

Acceptance - Rejection Technique [Thinning]

* Useful particularly when inverse cdf does not exist in closed form.
* Also called Thinning.
* To generate Random Variates x, uniformly distributed b/w $1/4$ and 1 ie $x \sim U(1/4, 1)$

Step 1:
Generate a Random Number $R \sim U(0,1)$

Step 2A:
If $R \geq 1/4$, Accept $x = R$. go to step 3

Step 2B:
If $R < 1/4$, Reject R, Return to step 1.

Step 3:
If another uniform Random Variate on $(1/4, 1)$ is needed, repeat the procedure beginning at step 1, if not STOP.

* R itself does not have desired distribution, but R conditioned on the even $(R \geq 1/4)$ does have the desired distribution.

proof:
Take $1/4 \leq a < b < 1$ Then

$$P\left(a < R \leq b \mid 1/4 \leq R \leq 1\right) = \frac{P(a < R \leq b)}{P(1/4 \leq R \leq 1)}$$

$$= \frac{b-a}{3/4}$$

which is the correct probability for a uniform distribution on $(1/4, 1)$

* <u>Efficiency:</u>

> Depends heavily on the ability to minimize the number of rejections.

→ Here the probability of rejection is $P(R < 1/4)$

So, $P(\text{success}) = p = \underline{3/4}$

$P(\text{failure}) = 1/p - 1 = \underline{1/3}\ (4/3 - 1)$

↳ P of rejection.

→ Mean no. of random numbers R required to generate one random variate x is one more than no. of rejections

Hence it is $4/3 - 1 + 1 = 1.33$

i.e, To generate 1000 values of x, would require approx 1333 random nos R.

Poisson Distribution

A poisson Random Variable N with mean $\alpha > 0$ has pmf,

$$p(n) = P(N = n) = \frac{e^{-\alpha} \cdot \alpha^n}{n!}, \quad n = 0, 1, 2 \ldots$$

$N \to$ no. of arrivals in one unit time.

procedure for generating a poisson random variate N -

<u>step 1</u>: set $n = 0$, $P = 1$.

<u>Step 2</u>: Generate a Random no. R_{n+1} and replace P by $P \cdot R_{n+1}$.

<u>Step 3</u>: if $p < e^{-\alpha}$, then accept $N = n$. otherwise, Reject increase n by one & return to step 2.

Problem:

Generate three poisson variates with mean $\alpha = 0.2$

Solution:

First, compute $e^{-\alpha}$.

$$e^{-\alpha} = e^{-0.2} = 0.8187$$

step 1: set $n=0$, $P=1$

step 2: $R_1 = 0.4357$ (Randomly taken)

$P_1 = 1 \cdot R_1$
$= 0.4357$

step 3: Since $P_1 < e^{-\alpha}$, accept $N=0$.

step 1: $n=0$, $P=1$

Step 2: $R_1 = 0.4146$

step 3: $P_1 < e^{-\alpha}$, accept $N=0$.

step 1: $n=0$, $P=1$

step 2: $R_1 = 0.8353$. $P = P * R_1 = 0.8353$

step 3: Since $P \geq e^{-\alpha}$, reject $N=0$, Goto step 2 with $n=1$.

step 2: $R_2 = 0.9952$, $P = P \cdot R_2 = 0.8353 * 0.9952$
$P = 0.8313$

Step 3: $P \geq e^{-\alpha}$. reject $N=1$, Goto step 2 with $n=2$

step 2: $R_3 = 0.8004$, $P = R_1 R_2 R_3 = 0.6654$.

step 3: Since $P < e^{-\alpha}$. accept $N=2$

Summary.

n	R_{n+1}	P	Accept / Reject	Result
0	0.4357	0.4357	Accept	$\alpha = 0$
0	0.4146	0.4146	Accept	$N = 0$
0	0.8353	0.8353	Reject	
1	0.9952	0.8313	Reject	
2	0.8804	0.6654	Accept	$N = 2$

Gamma Distribution

Step 1: ~~Gener~~

Compute $a = \frac{1}{(2\beta - 1)^{1/2}}$

$b = \beta - \ln 4$.

step 2:

Generate R_1 and R_2.

Set $V = \frac{R_1}{1 - R_1}$

step 3:

Compute $X = \beta V^a$

Step 4 A:

If $X > b + (\beta a + 1)\ln V - \ln(R_1^2 \cdot R_2)$.

Reject X and return to step 2.

step 4B:

~~else~~, use x as desired variate.

Step 5:

Replace X by $X/(\beta\theta)$.

SPECIAL PROPERTIES

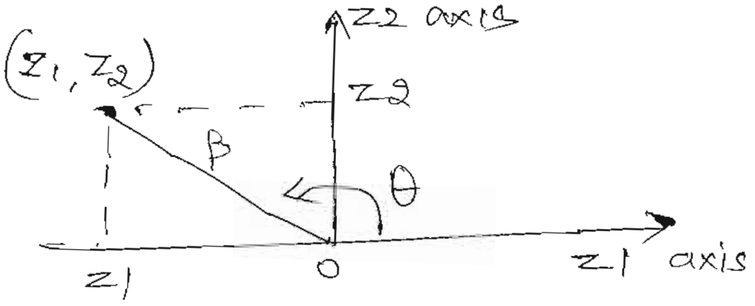
* "special properties" implies –
  They are variate generate techniques that are based on features of a particular family of probability distributions, rather than being general purpose techniques like inverse transform or A/R techniques.

* Here we see –
  1. Direct transformation for the Normal & logNormal distributions
  2. Convolution method.

Direct Transformation for the Normal and LogNormal Distributions.

a.) Approach for Normal (0, 1)

* Consider two standard normal random variables z_1 and z_2 plotted as a point in a plane



$$\left.\begin{aligned} z_1 &= B\cos\theta \\ z_2 &= B\sin\theta \end{aligned}\right\} \text{Polar Coordinates} \longrightarrow (1)$$

* WKT, $B^2 = z_1^2 + z_2^2$, has the chi-square distribution with 2 degrees of freedom, which is equivalent to an exp. distribution with mean 2

Hence,

$$B = \left(-2 \ln R\right)^{1/2} \longrightarrow (2)$$

* From (1) and (2), it gives a direct method for generating two independent standard normal variates z_1 and z_2, from two independent random nos R_1 and R_2

$$\boxed{\begin{aligned} z_1 &= (-2\ln R_1)^{1/2} \cos(2\pi R_2) \\ z_2 &= (-2\ln R_1)^{1/2} \sin(2\pi R_2) \end{aligned}}$$

b) Approach for normal (μ, σ^2)

Generate $z_i \sim N(0,1)$

$$\boxed{X_i = \mu + \sigma z_i}$$

(c) Approach for lognormal (μ, σ^2)

generate $X \sim N(\mu, \sigma^2)$

$$\boxed{Y_i = e^{X_i}}$$

Convolution Method

* The probability distribution of a sum of two or more independent random variables is called a convolution of the distributions of the original variables.

* **Erlang Distribution.**

An erlang ~~dist.~~ random variable X with parameters (K, θ) can be shown to be the sum of K independent exponential random variables

$X_i \ (i = 1, 2 \ldots K)$, each having a mean $1/K\theta$

$$X = \sum_{i=1}^{K} X_i$$

Erlang variate can be generated by:

$$X = \sum_{i=1}^{K} \frac{-1}{K\theta} \ln R_i$$

$$\boxed{X = \frac{-1}{K\theta} \ln \left(\prod_{i=1}^{K} R_i \right)}$$

ASHOK KUMAR. K
9742024066
celestialcluster@gmail.com.

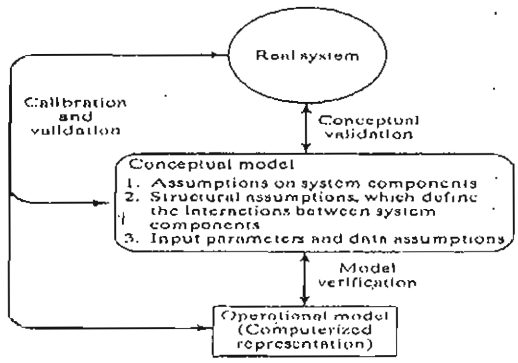
UNIT 8:

VERIFICATION AND VALIDATION OF SIMULATION MODELS, OPTIMIZATION

- The goal of the validation process is:
  - To produce a model that represents true behavior closely enough for decision-making purposes
  - To increase the model's credibility to an acceptable level
- Validation is an integral part of model development
- **Verification:** Building the model correctly (correctly implemented with good input and structure). It proceeds by the comparisons of the conceptual model to the computer representation that implements the conception.
  Validation: Building the correct model (an accurate representation of the real system). It is usually achieved through the calibration of the model.

MODEL BUILDING, VERIFICATION, AND VALIDATION

- The first step in model building consists of observing the real system and the interactions among its various components and collecting data on its behavior. Operators, technicians, repair and maintenance personnel, engineers, supervisors, and managers understand certain aspects of the system that may be unfamiliar to others. As model development proceeds, new questions may arise, and the model developers will return to this step of learning true system structure and behavior.
- The second step in model building is the construction of a conceptual model--a collection of assumptions on the components and the structure of the system, plus hypotheses on the values of model input parameters, illustrated by the following figure.
- The third step is the implementation of an operational model. i.e, translation of the operational model into a computer recognizable form- the computerized model.



VERIFICATION OF SIMULATION MODELS

- **Purpose:** Assure that the conceptual model is reflected accurately in the operational model.
- The conceptual model quite often involves some degree of abstraction about system operations, or some amount of simplification of actual operations.
- Many common-sense suggestions can be given for use in the verification process:
  - Have the operational model checked by someone other than its developer.
  - Make a flow diagram which includes each logically possible action a system can take when an event occurs, and follow the model logic for each action for each event type.
  - Closely examine the model output for reasonableness under a variety of settings of input parameters.
  - Have the operational model print the input parameters at the end of the simulation, to be sure that these parameter values have not been changed inadvertently.
  - Make the operational model as self-documenting as possible.
  - If the operational model is animated, verify that what is seen in the animation imitates the actual system.
  - The interactive run controller (IRC) or debugger is an essential component of successful simulation model building. Even the best of simulation analysts makes mistakes or commits logical errors when building a model. The IRC assists in finding and correcting those errors in the follow ways:
    (a) The simulation can be monitored as it progresses.
    (b) Attention can be focused on a particular line of logic or multiple lines of logic that constitute a procedure or a particular entity.
    (c) Values of selected model components can be observed. When the simulation has paused, the current value or status of variables, attributes, queues, resources, counters, etc., can be observed.
    (d) The simulation can be temporarily suspended, or paused, not only to view information but also to reassign values or redirect entities.
  - Graphical interfaces are recommended for accomplishing verification & validation

CALIBRATION AND VALIDATION OF MODELS

- Although verification and validation are conceptually distinct, they are usually conducted simultaneously by the modeler.

- **Validation:** The overall process of comparing the model and its behavior to the real system.

  Calibration: The iterative process of comparing the model to the real system, making adjustments to the model, comparing the revised model to reality, making additional adjustments, comparing again, and so on.

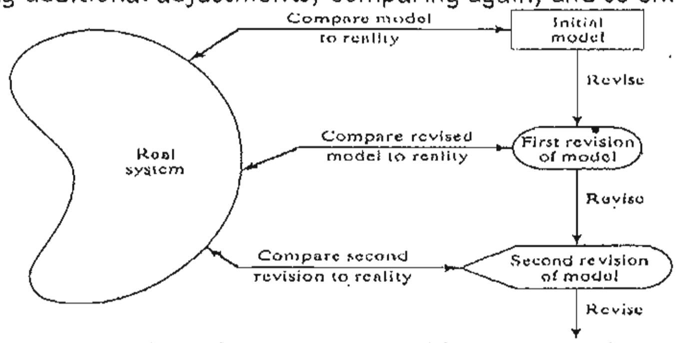


Fig: iterative process of calibrating a model.

- No model is ever a perfect representation of the system. The modeler must weigh the possible, but not guaranteed, increase in model accuracy versus the cost of increased validation effort.
- As an aid in the validation process, Naylor and Finger [1967] formulated a three step approach which has been widely followed:
  - Build a model that has high face validity.
  - Validate model assumptions.
  - Compare the model input-output transformations to corresponding input-output transformations for the real system.

The next five subsections investigate these three steps in detail.

Face Validity

- The first goal of the simulation modeler is to construct a model that appears reasonable on its face to model users and others who are knowledgeable about the real system being simulated.
- The potential users of a model should be involved in model construction from its conceptualization to its implementation, to ensure that a high degree of realism is built into the model through reasonable assumptions regarding system structure, and reliable data.
- Sensitivity analysis can also be used to check a model's face validity.

Validation of model assumptions

- Model assumptions fall into two general classes:
  - **Structural assumptions:** involve questions of how the system operates and usually involves simplifications and abstractions of reality.

- **Data assumptions:** based on collection of reliable data and its correct statistical analysis

- **Bank example:** customer queuing and service facility in a bank.
  Structural assumptions, e.g., customer waiting in one line versus many lines, served FCFS versus priority.
  Data assumptions, e.g., interarrival time of customers, service times for commercial accounts.
- The procedure for analyzing input data consist of three steps:-
  1: Identifying the appropriate probability distribution.
  2: Estimating the parameters of the hypothesized distribution .
  3: Validating the assumed statistical model by goodness - of - fit test such as
  the chi-square test, KS test and by graphical methods.

Validating input-output transformations

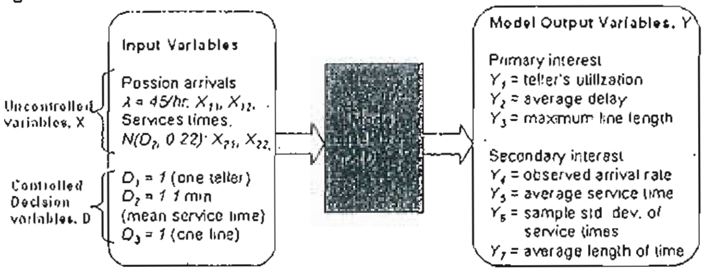
1. The ultimate test of a model, and in fact the only objective test of the model as a whole, is the model's ability to predict the future behavior of the real system when the model input data match the real inputs and when a policy implemented in the model is implemented at some point in the system.
2. In this phase of validation process, the model is viewed as input-output transformation--that is, the model accepts values of the input parameters and transforms these inputs into output measures of performance. It is this correspondence that is being validated.
3. Instead of validating the model input-output transformation by predicting the future, the modeler could use past historical data which has been reserved for validation purposes only--that is, if one data set has been used to develop and calibrate the model, it is recommended that a separate data set be used as the final validation test.
4. Thus accurate "prediction of the past" may replace prediction of the future for the purpose of validating the model.
5. A necessary condition for the validation of input-output transformations is that some versions of the system under study exist, so that the system data under at least one set of input conditions can be collected to compare to model predictions.
6. If the system is in planning stage and no system operating data can be collected, complete input-output validation is not possible.
7. Validation increases modeler's confidence that the model of existing system is accurate.
8. Changes in the operational model, ranging from relatively minor to relatively major include:
   1: Minor changes of single numerical parameters, such as speed of the machine, arrival rate of the customer etc.
   2: Minor changes of the form of a statistical distribution, such as distribution of service time or a time to failure of a machine.

3: Major changes in the logical structure of a subsystem, such as change in queue discipline for waiting-line model, or a change in the scheduling rule for a job shop model.

4: Major changes involving a different design for the new system, such as computerized inventory control system replacing a non computerized system

Bank Example:

- One drive-in window serviced by one teller, only one or two transactions are allowed.
  - Data collection: 90 customers during 11 am to 1 pm.
    Observed service times $\{S_i, i = 1,2, \ldots, 90\}$.
    Observed interarrival times $\{A_i, i = 1,2, \ldots, 90\}$.
  - Data analysis let to the conclusion that:
    Interarrival times: exponentially distributed with rate $l = 45$
    Service times: $N(1.1, 0.22)$
- A model was developed in close consultation with bank management and employees.
- Model assumptions were validated.
- Resulting model is now viewed as a "black box":



- Comparison with Real System Data
  - Real system data are necessary for validation.
    - o System responses should have been collected during the same time period (from *11*am to *1*pm on the same Friday.)
  - Compare the average delay from the model $Y2$ with the actual delay $Z2$:
    - o Average delay observed, $Z2 = 4.3$ minutes, consider this to be the true mean value $m0 = 4.3$.
    - o When the model is run with generated random variates $X1n$ and $X2n$, $Y2$ should be close to $Z2$.
    - o Six statistically independent replications of the model, each of 2-hour duration, are run.
    - o Null hypothesis testing: evaluate whether the simulation and the real system are *the same* (w.r.t. output measures):

$$H_0: E(Y_2) = 4.3 \text{ minutes}$$

$$H_1: E(Y_2) \neq 4.3 \text{ minutes}$$

ASHOK KUMAR K

- If $H0$ is not rejected, then, there is no reason to consider the model invalid
- If $H0$ is rejected, the current version of the model is rejected, and the modeler needs to improve the model

- Conduct the t test:
  - Chose level of significance ($\alpha = 0.05$) and sample size ($n = 6$), see result in Table 10.2.
  - Compute the same mean and sample standard deviation over the n replications:

$$\bar{Y}_2 = \frac{1}{n}\sum_{i=1}^{n} Y_{2i} = 2.51 \text{ minutes} \qquad S = \sqrt{\frac{\sum_{i=1}^{n}(Y_{2i} - \bar{Y}_2)^2}{n-1}} = 0.81 \text{ minutes}$$

  - Compute test statistics:

$$|t_0| = \left|\frac{\bar{Y}_2 - \mu_0}{S/\sqrt{n}}\right| = \left|\frac{2.51 - 4.3}{0.82/\sqrt{6}}\right| = 5.24 > t_{critical} = 2.571 \text{ (for a 2-sided test)}$$

  - Hence, reject H0. Conclude that the model is inadequate.
  - Check: the assumptions justifying a t test, that the observations (Y2i) are normally and independently distributed.

- Similarly, compare the model output with the observed output for other measures: $Y4 \leftrightarrow Z4$, $Y5 \leftrightarrow Z5$, and $Y6 \leftrightarrow Z6$

Input-output validation: using historical input data

- When using artificially generated data as input data, the modeler expects the model to produce event patterns that are compatible with, but not identical to, the event patterns that occurred in the real system during the period of data collection.
- Thus, in the bank model, artificial input data {X\n, X2n, n = 1,2, , .} for inter arrival and service times were generated and replicates of the output data Y 2 were compared to what was observed in the real system
- An alternative to generating input data is to use the actual historical record, {An, Sn, n = 1,2,...}, to drive simulation model and then to compare model output to system data.
- To implement this technique for the bank model, the data Ai, A2,..., S1 S2 would have to be entered into the model into arrays, or stored on a file to be read as the need arose.

- To conduct a validation test using historical input data, it is important that all input data (An, Sn,...) and all the system response data, such as average delay(Z2), be collected during the same time period.

Input-output validation: using a Turing test

- Used in addition to statistical test, or when no statistical test is readily applicable. It utilizes persons' knowledge about the system to compare model output to the system output.
- For example:
  - Present 10 system performance reports to a manager of the system. Five of them are from the real system and the rest are "fake" reports based on simulation output data.
  - If the person identifies a substantial number of the fake reports, interview the person to get information for model improvement.
  - If the person cannot distinguish between fake and real reports with consistency, conclude that the test gives no evidence of model inadequacy.

Topic left:
optimization via simulation

Materials handling sim
liquified natural gas transp
Automobile Engine assembly
Sequencing Traffic Signal
Online Services

UNIT 1: INTRODUCTION

1. What is simulation? Explain with flowchart, the steps involved in simulation study.

   - May 2010, 10 Marks
   - Dec 2011, 10 Marks.
   - June 2011, 10 Marks.

2. Explain when simulation is appropriate tool and not.

   - Dec 2011, 10 Marks.

3. Explain the advantages and disadvantages of simulation.

4. What is System and System Environment? List the Components of the System with example.

   - May 2010, 5 Marks
   - June 2011, 10 Marks

5. Differentiate between continuous and discrete systems

   - May 2010, 05 marks

6. What is a Model of the system? Explain the different types of models.

7. With the help of flow diagram, explain the simulation of single channel queuing system

   - June 2011, 10 Marks.

8. Describe the different areas of applications of simulation.

a.) <u>problem Formulation</u>

→ In this step, the problem is clearly stated.
→ It can be either provided by policy makers or analyst.

b.) <u>Setting of objectives and overall project plan.</u>

→ In this step, we decide how we should approach the problem
→ A statement of alternative systems also considered.

c.) <u>Model conceptualization</u>

→ In this step, we will establish a reasonable model.
→ Model building is an art learnt over time and experience.

d.) <u>Data Collection</u>

→ Here, we will collect the data necessary to run the simulation.
ex:- Arrival rate, Service rate etc

e.) <u>Model Translation</u>

→ Here we will convert a model into a programming language, because real systems
- Involves lot of information storage
- Involves lot of computation.

f.) <u>Verified ?</u>

→ verify the model by checking if the program works properly.

g.) <u>Validated ?</u>

→ check if the system accurately represents the real system.

1. What is Simulation? Explain with flowchart the steps involved in simulation study.

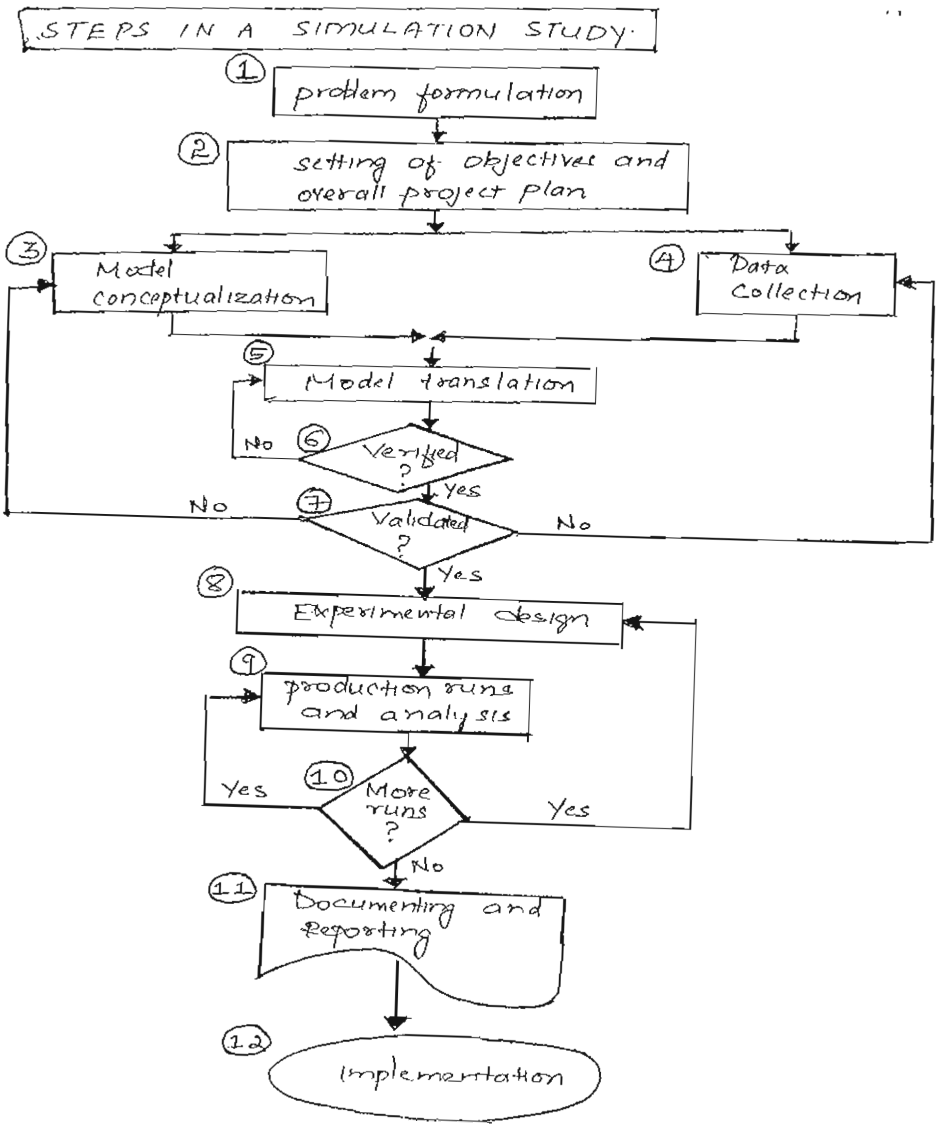
* Simulation

→ Simulation is the imitation of the operation of a real-world process or system over time.

→ Simulation modeling can be used

- As an analysis tool for predicting the effect of changes to existing system.
- As a design tool to predict the performance of new system.

* steps involved in a simulation study



h.) <u>Experimental Design</u>

→ How many runs?
→ How long?
→ What kind of input variation?

} Design is made on all these.

i.) <u>production runs and Analysis</u>

→ Actual running of the simulation.
→ Collect and analyze the output.

j.) <u>More Runs?</u>

→ Repeat the experiment, if necessary.

k.) <u>Documentation and Reporting</u>

→ Document and report the results.
→ There are two types of documentations
- program documentation
- progress documentation

l.) <u>Implementation</u>

→ The model is implemented practically depending on the success/failure of all the above steps.

2. Explain when simulation is appropriate tool and when it is not.

When simulation is a Appropriate tool ?

a.) Simulation enables the study of, and experimentation with internal interactions of a complex system.

b.) Informational, organizational, and environmental changes can be simulated, and the effect of these alterations on model behavior can be observed.

c.) The knowledge gained in designing a simulation model may be of great value towards suggesting improvement in the system under investigation.

d.) It helps in understanding which variables are most important and how variables interact.

e.) It can be used as a pedagogical device to reinforce analytic solutions.

f.) It can be used to verify analytic solutions.

g.) By simulating different capabilities for a machine, requirements can be determined.

h.) Simulation models designed for training allows learning without the cost.

i.) Animation shows a system in simulated operation so that the plan can be visualized.

j.) The modern system is so complex that the interactions can be treated only through simulation.

When simulation is NOT Appropriate?

a.) When the problem can be solved using common sense

b.) when the problem can be solved Analytically.

c.) When it is easier to perform Direct Experiments.

d.) When the simulation cost exceeds savings.

e.) when resources or time are not available.

f.) when the system behavior is too complex.

g.) when there isn't the ability to verify and validate the model.

3. Explain the Advantages and Disadvantages of simulation.

Advantages of simulation

i.) New policies, operating procedures, decison rules, organizational procedures and so on can explored without disrupting the ongoing operations.

.) New hardware designs, physical layouts, transportation systems, and so on can be tested without committing resources for their acquisition.

:) Hypothesis about how and why certain phenomenon occurs can be tested for feasibility.

d.) Insight can be obtained about the interaction of variables.

e.) Insight can be otained about importance of variables

f.) Bottleneck analysis can be performed indicating work in process, information and so on are being excessively delayed.

g.) Simulation study can help in understanding how the system operates rather than how the individuals think the system operates.

h.) "What-if" questions can be answered. This is useful in designing the new system.

Disadvantages of simulation

a.) Model building requires <u>special training</u>.

→ It is an art that is learnt over time and experience

b.) Simulation results may be <u>difficult to interpret</u>.

→ Since most simulation outputs are essentially random variables, it may be hard to determine whether an observation is a result of system interrelationships or randomness.

c.) Simulation Modeling and Analysis can be <u>time consuming and expensive</u>.

→ Skimping on resources for modelling and analysis may result in a simulation model that is not sufficient for the task.

d.) Simulation is used in some cases when an <u>analytic solution is possible or even preferrable</u>.

→ This may be particularly true in the simulation of some waiting lines.

4. What is System and System Environment? List the components of the system with Examples

System

* A system is defined as a group of objects that are joined together in some regular interaction or interdependent toward the accomplishment of some purpose.

System Environment.

* A system is often affected by changes occuring outside the system. Such changes are said to occur in system Environment.

Components of the system.

* Entity
  → An object of interest in the system
  → Ex: customers, machines.

* Attributes
  → A property of an entity
  → Ex: checking account balance.

* Activities
  → Time period of specified length
  → ~~Mak~~ Ex: Make Deposits.

* State
  → The Collection of variables necessary to describe the system at any time, relative to the objectives of the study.

* **Event**

→ An instantaneous occurrence that may change the state of the system.

→ Ex: Arrivals, Departures.

* **Endogeneous**

→ Used to describe activities and events occuring within a system.

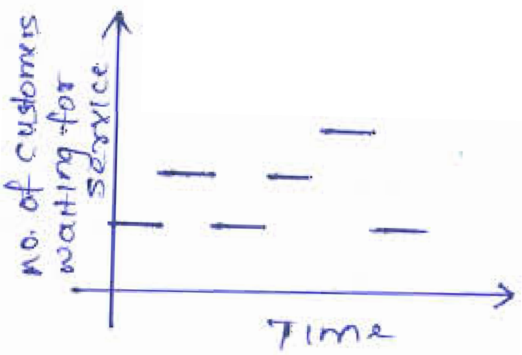
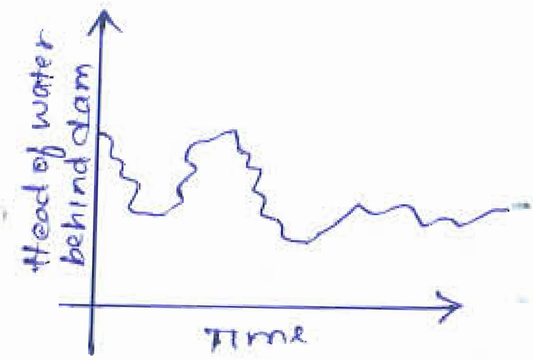
* **Exogeneous**

→ Used to describe activities and events in an environment that affects the system.

Examples of system and components.

System	BANKING	PRODUCTION
Entity	Customers	Machines
Attributes	checking Balance	Speed, capacity
Activities	Make Deposits	Welding, stamping
Events	Arrival, Departure	Breakdown
State variables	No. of cust. waiting, No. of Busy Tellers	status of m/c [Busy, Idle, Down]

5. Differentiate between continuous and Discrete Systems.

Discrete System	Continuous System
* A Discrete System is one in which the state variables change only at a discrete set of points	* A continuous system is one in which the state variables change continuously over time
* 	* 
* Example: A Bank	* Example: Head of water behind a dam.

UNIT 2: GENERAL PRINCIPLES, SIMULATION SOFTWARES.

1. Briefly describe the concepts used in discrete event simulation with examples.

   – May/June 2010

2. Explain Event scheduling Algorithm by generating system snapshots at clock = t1 and clock = t

   – May/June 2010, 6 Marks.

3. What do you mean by world view? Discuss the various types of world views.

   – June/July 2011, 10 Marks.

4. What is List processing? Explain the basic properties and operations on lists

5. Explain simulation in Java, with a block diagram, for single server queue simulation.

   – December 2011, 6 marks.

6. Explain simulation in GPSS, with a neat block diagram for single server queue simulation.

ASHOK KUMAR. K
celestialcluster@gmail.com
9742024066

6. What is a Model of a system ? Explain the different types of models.

Model

* Model is a representation of a system for the purpose of studying the system.
* It is the simplification of the system.

Types of Model

a.) Static v/s Dynamic

b.) Deterministic v/s Stochastic

Static v/s Dynamic Simulation models

* Static simulation model (Monte Carlo Simulation) represents a system at a particular point in time.
* Dynamic simulation model represents the system as they change over time.

Deterministic v/s Stochastic Simulation models.

* Deterministic simulation model contains no random variables and have a known set of inputs which will result in a unique set of outputs.
* Stochastic Simulation model has one or more random variables as inputs. Random inputs leads to random outputs.

7. With the help of flow diagram, explain the simulation of single channel queuing system.

Queuing system

* A queuing system is described by its calling population, the nature of arrivals, the service mechanism, system capacity, and queuing discipline.

Simulation of single channel queuing system.

* Here the calling population is infinite
* Arrivals for service occurs one at a time in a random fashion.
* Service times are of some random length according to a probability distribution which does not change over time.
* The system capacity has no limit, meaning that any number of units can wait in line.
* Finally, the units are served in the order of their arrival (FCFS) by a single server.
* Arrivals and services are defined by some probability distributions.

* **Definitions**

→ **System State**: The number of units in the system and the status of the server (busy or idle)

→ **Event**: A set of circumstances that cause an instantaneous change in the state of the system. In single channel queuing system, there are only two possible events -

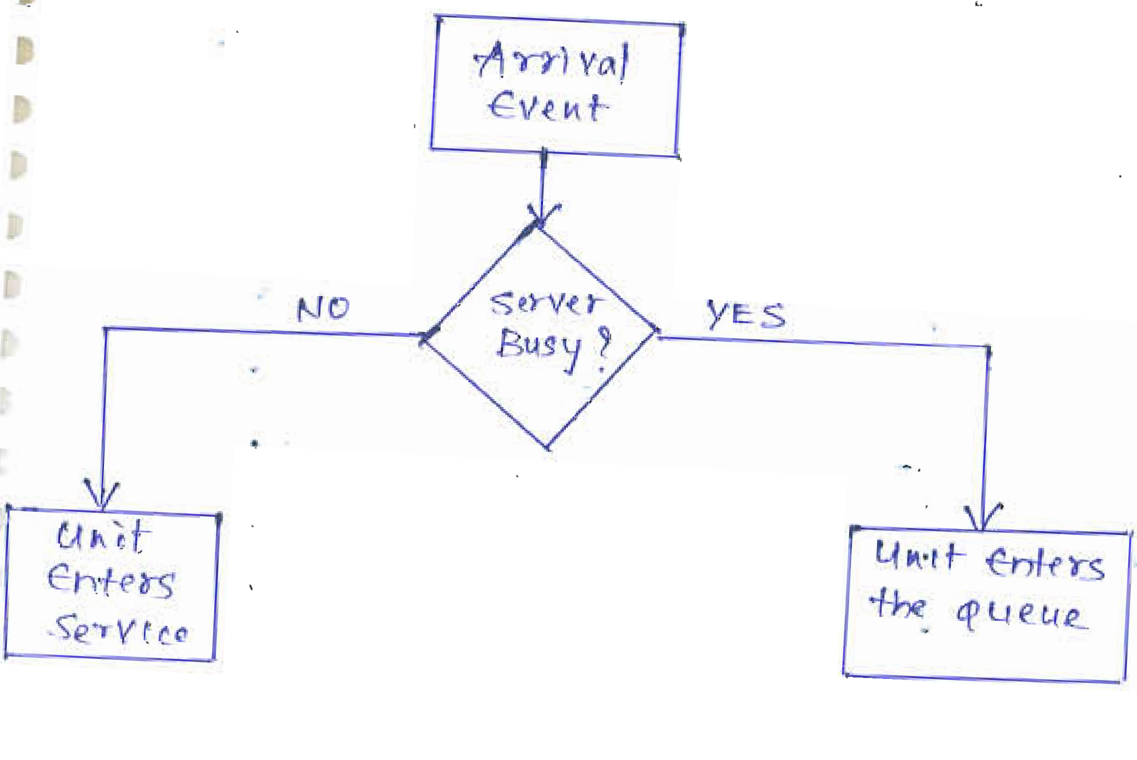
i.) **Arrival Event**: Entry of a unit into system.

ii.) **Departure Event**: Completion of service on a unit

→ **clock**: used to track simulated time

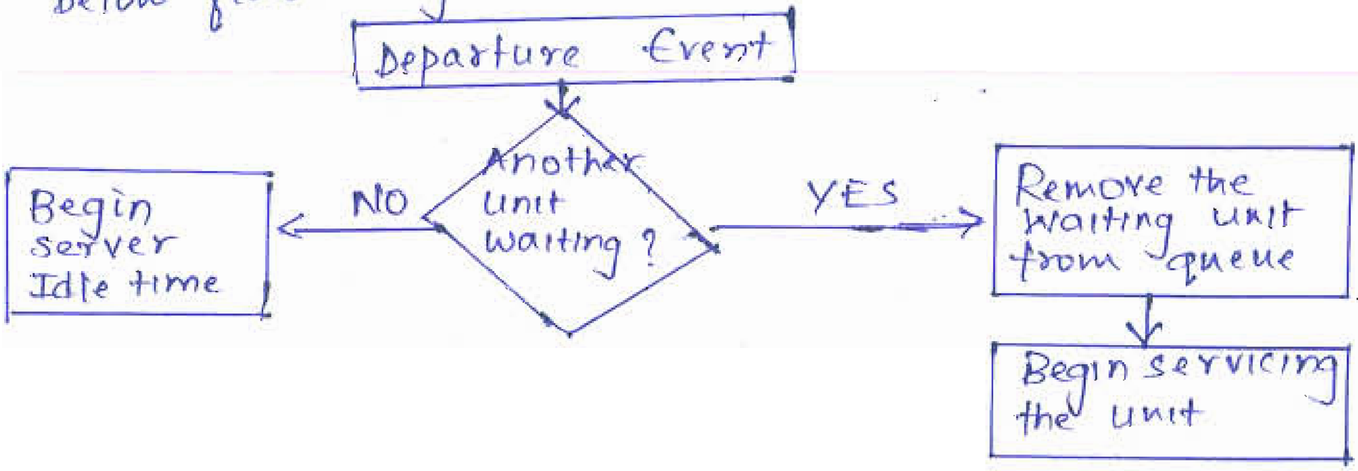
Arrival Event

* If a unit has just entered the system, the simulation proceeds in the manner shown in the below flow diagram.



Departure Event

- If a unit has just completed the service, the simulation proceeds in the manner shown in the below flow diagram.



- Fig below shows potential unit actions upon arrival.

		Queue Status	
		Not empty	Empty
Server Status	Busy	Enter Queue	Enter queue
	Idle	Impossible	Enter service

- Fig below shows server outcomes after the completion of service.

		Queue Status	
		Not Empty	Empty
Server status	Busy		Impossible
	Idle	Impossible	

7. Describe the different areas of applications of simulation

a.) Manufacturing Applications

b.) Semiconductor manufacturing

c.) Construction Engineering and project management

d.) Military Applications.

e.) Logistics, supply chain, and distribution Applications.

f.) Transportation modes and Traffic.

g.) Business process simulation.

h.) Health care.

ASHOK KUMAR. K
celestialcluster@gmail.com
9742024066

i.) Event list
list of event notices for future events, ordered by time of occurrence.
Also known as Future event list.

j.) Activity
A duration of time of specified length.
eg: service time

k.) Delay
A duration of time of unspecified length.
eg: customer's waiting time in queue.

l.) clock
A variable representing simulated time.

Examples (Able - Baker)

→ System state
- $L_Q(t)$ → no. of cars waiting in the queue at time t.
- $L_A(t)$ → no. of cars servicing by Able at time t
- $L_B(t)$ → no. of cars servicing by Baker at time t.

→ Entities
cars, servers.

→ Events
- Arrival Event
- service completion by Able
- service completion by Baker

→ Activities
- Inter arrival time
- service time by Able
- service time by baker

→ Delay
A car's wait in queue until Able or Baker becomes free.

. **Briefly describe the concepts in discrete event simulation with examples.**

Concepts in Discrete Event Simulation

1.) **System**
Collection of entities that interact together over time to accomplish one or more goals.

b.) **Model**
An abstract representation of a system, usually containing structural, logical, or mathematical relationships which describe the system.

c.) **System state**
A collection of variables that contain all the information necessary to describe the system at any time.

d.) **Entity**
Any object or component in the system which requires explicit representation in the model.

e.) **Attributes**
The properties of a given entity.

f.) **List**
Collection of associated entities ordered in some logical fashion.

g.) **Event**
An instantaneous occurrence that changes the state of a system.

h.) **Event notice**
A record of an event to occur at the current or some future time, along with any associated data necessary to execute the event.

2. **Explain Event scheduling Algorithm by generating system snapshots at clock = t and clock = t1**

Event scheduling / Time Advance Algorithm.

* After the system snapshot at CLOCK = t has been updated, the CLOCK is advanced to the imminent simulation time CLOCK = t1. The imminent event notice is removed from FEL, and the event is executed.
* This process repeats until the simulation is over.
* The removal and addition of events from FEL illustrated in below Figure:

old system snapshot at time t

CLOCK	System state	....	F.E.L	....
t	(5,1,6)		(3,t1) → type 3, time t1 (1,t2) (1,t3) ⋮ (2,tn)	

Event scheduling / time advance Algorithm

Step 1: Remove the event notice for the imminent event (event 3, time t1) from FEL.

Step 2: Advance the clock to imminent event time. (ie Advance the CLOCK from t to t1)

Step 3: Execute the imminent event; update system state. change entity attributes, and set memberships as needed.

Step 4: Generates future events (if necessary) and place their event notices on FEL ranked by event time.

Ex: event 4 to occur at time t^* where $t_2 < t^* < t_3$.

Step 5: Update cumulative statistics and counters.

New system snapshot at time t_1

CLOCK	system state	....	F.E.L	....
t_1	(5, 1, 5)		$(1, t_2)$ → type 1, time t_2 $(4, t^*)$ $(1, t_3)$ ⋮ $(2, t_n)$	

- One possible way to determine the correct position of event notice on the FEL is to conduct a top-down search

if $t^* < t_2$ place event 4 at top of FEL

if $t_2 < t^* < t_3$ place event 4 second on the list

if $t_3 < t^* < t_4$ place event 4 third on the list.

⋮

if $t_n < t^*$, place event 4 last on the list.

· What do you mean by world view?
· Discuss the various types of world views.

World view

When using a simulation package or even when using a manual simulation, a modeler adopts a world view or orientation for developing a model.

Various types of world views

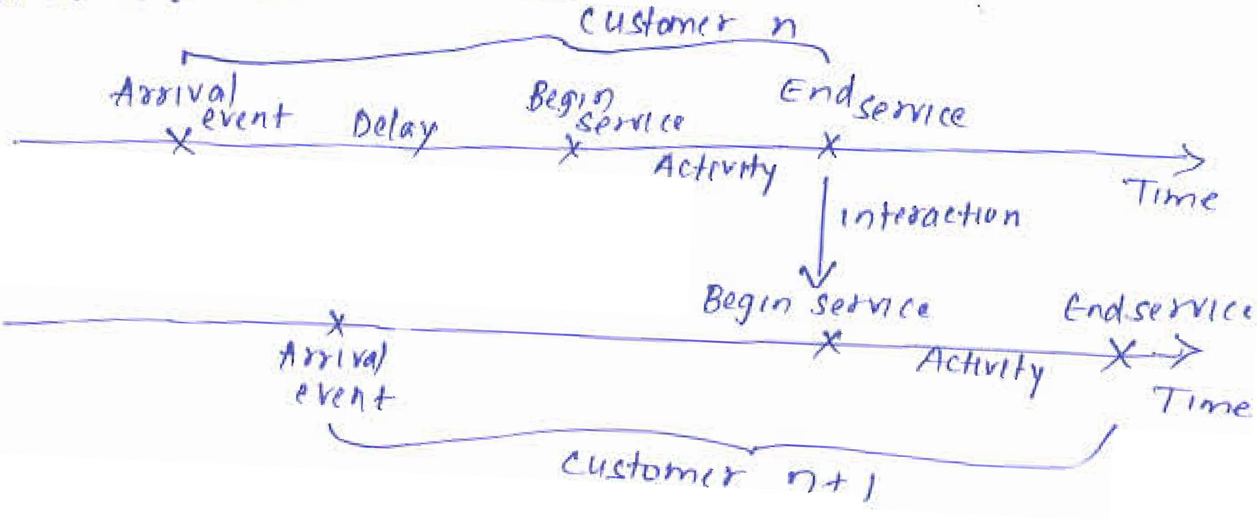
a.) Event scheduling world views
b.) process - Interaction world views
c.) Activity - scanning world views.

Event scheduling world view

* Here, a simulation analyst concentrates on events and their effect on system state.

process interaction world view

* Here, a simulation analyst thinks in terms of processes.
* A process is the life cycle of one entity. This life cycle consists of various events and activities.
* Fig shows the interaction b/w two customer processes as customer n+1 is delayed until the previous customer's "end-service" event occurs.



The main operations on a list are:

a.) Removing a record from the top of the list.

→ when time is advanced and the imminent event is due to be executed.

→ By adjusting the head pointer on the FEL ⇔ by removing the event at the top of FEL.

b.) Removing a record from any location on the list.

→ if an arbitrary event is being canceled, or an entity is removed from a list based on some of its attributes, to begin an activity.

→ By making a partial search through the list.

c.) Adding an entity record to the top or bottom of the list

→ When an entity joins the back of a FIFO queue

→ By adjusting the tail pointer on the FEL ⇔ by adding an entity to the bottom of the FEL

d.) Adding a record to an arbitrary position on the list determined by the ranking rule.

→ If a queue has a ranking rule of earliest due date first (EOF)

→ By making a partial search through the list.

Explain the simulation in Java, with a block diagram for a single server queue simulation.

- Java has been used extensively in simulation.
- It does not provide any facilities directly, so the simulation analyst must program all details of event-scheduling / time advance algorithm, the statistic-gathering capabilities etc, however the runtime library does provide a random number generator.
- the following components are common to almost all models written in Java:

CLOCK: A variable defining simulated time.

Initialization method:
A method to define the system state at time 0.

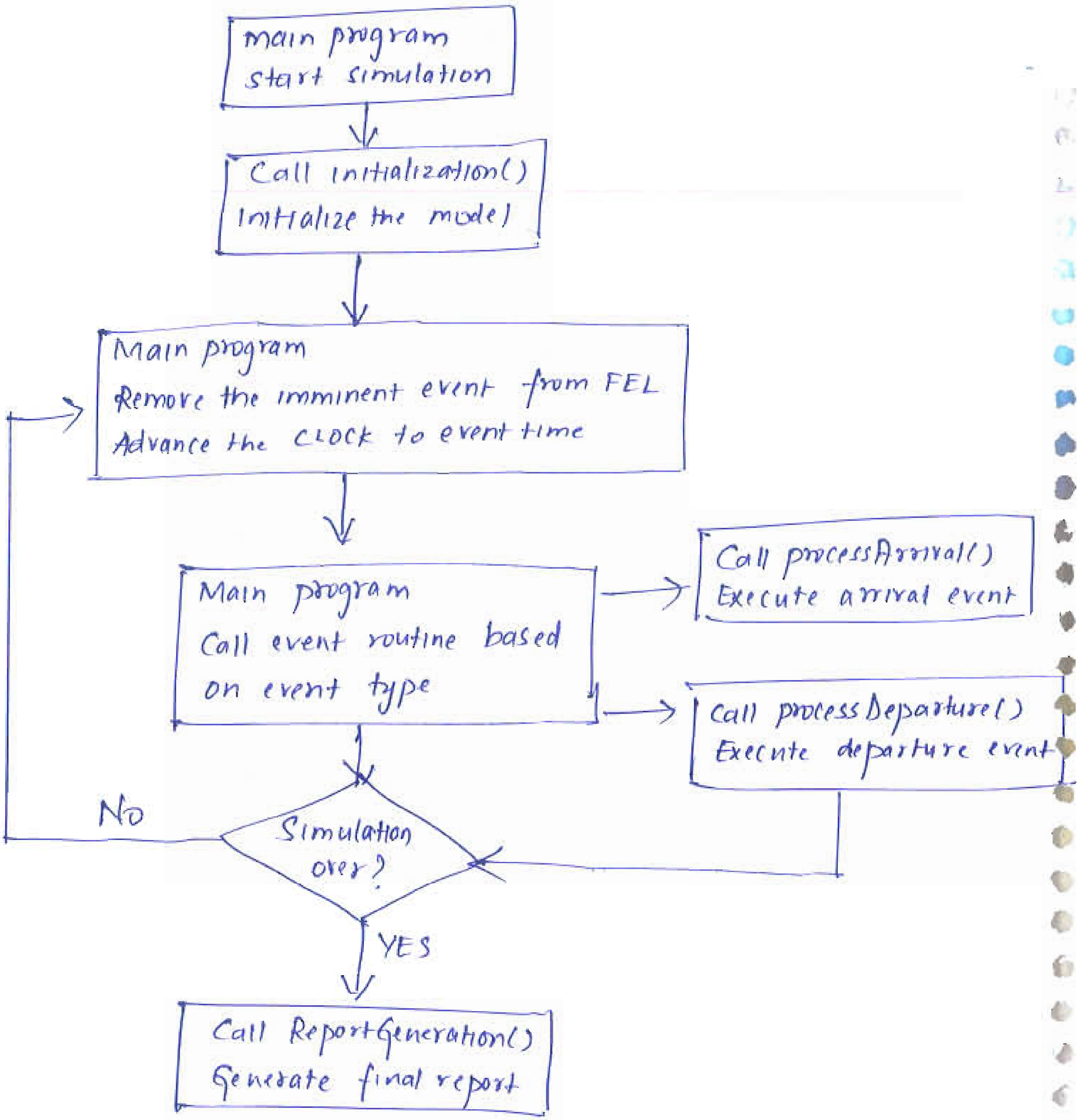
Min-time Event method:
It identifies the imminent events that has smallest time stamp.

Event method:
For each type, a method to update system state when that event occurs.

Random variate Generators:
Methods to generate samples from desired probability distribution.

Main program
To maintain overall control of the event scheduling algorithm.

Report Generator
Computes summary from cumulative statistics and prints the report.



* Above fig shows over all structure of Java simulation of a single server queue.

c. Explain the simulation in GPSS, with neat diagram for single channel queue simulation.

Simulation in GPSS

* GPSS is highly structured, special purpose simulation programming language based on the process-interaction approach and oriented toward queuing systems.

* There are over 40 standard blocks in GPSS. Entities called transactions may be viewed as flowing through the block diagram. Blocks represent events, delays, and other actions that affect transaction flow.

* Thus, GPSS can be used to model any situation where transactions (entities, customers, units of traffic) are flowing through a system. The block diagram is converted into block statements, control statements are added, and the result is a GPSS model.

* IBM (International Business Machines) was the first to release the initial version of GPSS.

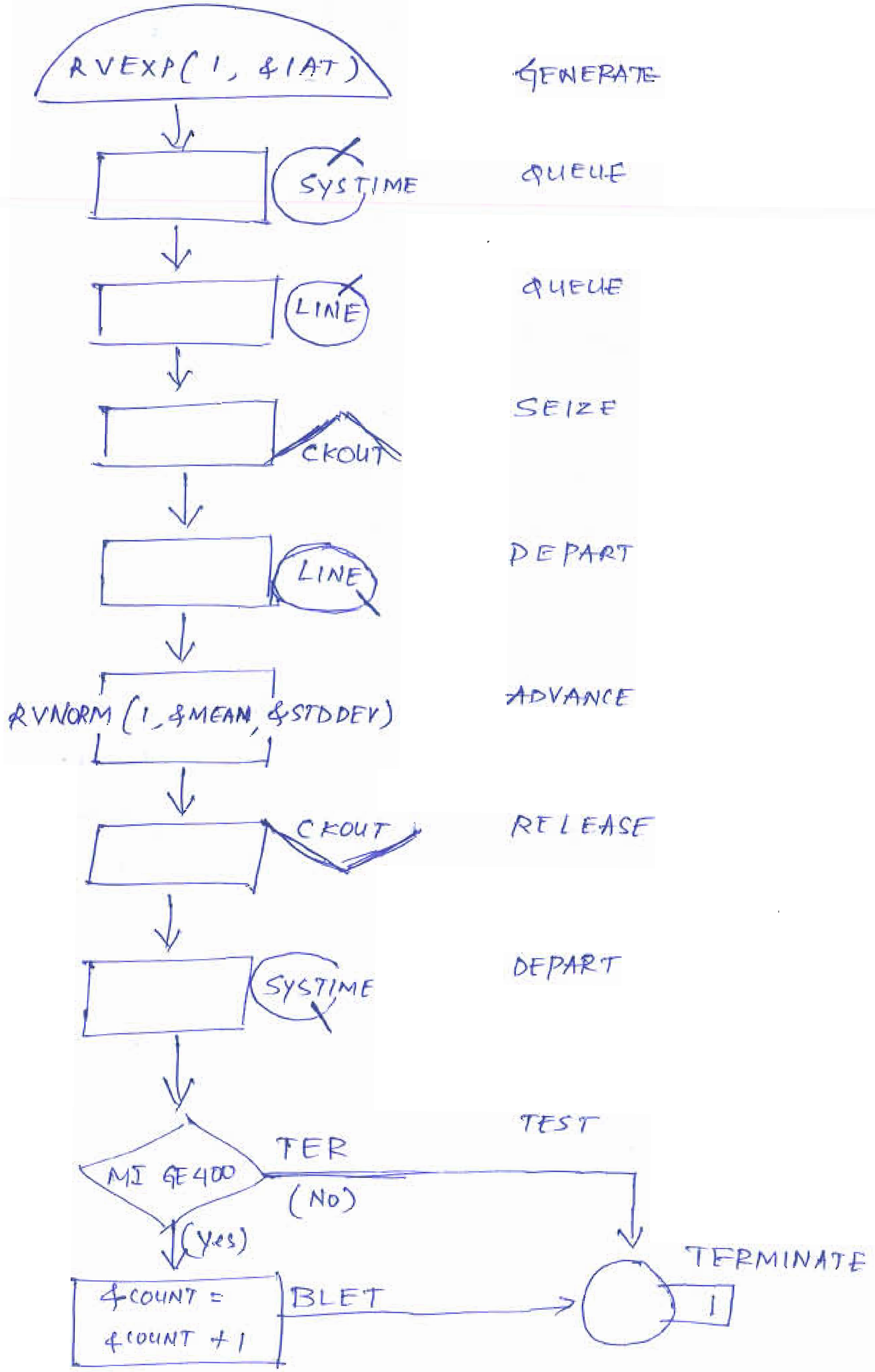


Fig: GPSS block diagram for the single-server queue simulation

UNIT 3: STATISTICAL MODELS IN SIMULATION.

1. Explain the concepts and terminologies involved in statistical models in simulation.

   (or) Explain Discrete Random variables and continuous Random variables with Examples

   — Dec 2011, 10 Marks.

2. Explain the following Discrete Distributions
   - a.) Bernoulli process
   - b.) Binomial distributions
   - c.) Geometric distributions
   - d.) Negative Binomial distributions.

3. Explain the following continuous Distributions
   - a.) uniform Distributions
   - b.) Exponential Distributions.

4. Explain the poissons process in detail.

5. Explain the memoryless property of an exponential Distribution. prove it mathematically

   [refer quest. no. 3 for Answer]

6. Write a short note on Non Stationary poissons process

   [refer Quest no. 4 for Answer]

Ashok Kumar. K
9742024066
celestialcluster@gmail.com

1. Explain the concepts and Terminologies used in Statistical models.

Concepts and Terminologies in Statistical models

a.) Discrete Random Variables

b.) Continuous Random Variables

c.) Cumulative Distribution function

d.) Expectation and Variance

e.) Mode.

Discrete Random Variables

* Definition:

Let X be any Random variable. If the number of possible values of X is finite or countably infinite, X is called a Discrete Random Variable

Ex: X: No. of Heads in tossing a coin twice.

* Set of all possible values of X is defined as a Range space of X

$$R_X = \{0, 1, 2 \ldots\}$$

* $p(x_i) = P(X = x_i)$ gives the probability that the random variable equals the value x_i.

* $p(x_i)$ has to Satisfy following conditions

→ $p(x_i) \geq 0, \; \forall i$

→ $\sum_{i=1}^{\infty} p(x_i) = 1$

* Collection of pairs $(x_i, p(x_i))$ is called the probability distribution of X, and $p(x_i)$ is called probability mass function (pmf) of X.

Continuous Random Variables

* **Definition:**

If the Range space of X is an interval or collection of intervals, then the Random variable X is called a Continuous Random Variable

Ex: X: Spinning of a magnetic needle.

* probability of X lies in the interval $[a, b]$ is given by -

$$P(a \le x \le b) = \int_a^b f(x)\,dx \quad \text{——— (1)}$$

* The function $f(x)$ is called the probability density function (pdf) of X.

* PDF satisfies the following properties

→ $f(x) \geq 0 \quad \forall \; x \in R_X$

→ $\int_{R_X} f(x)\,dx = 1$

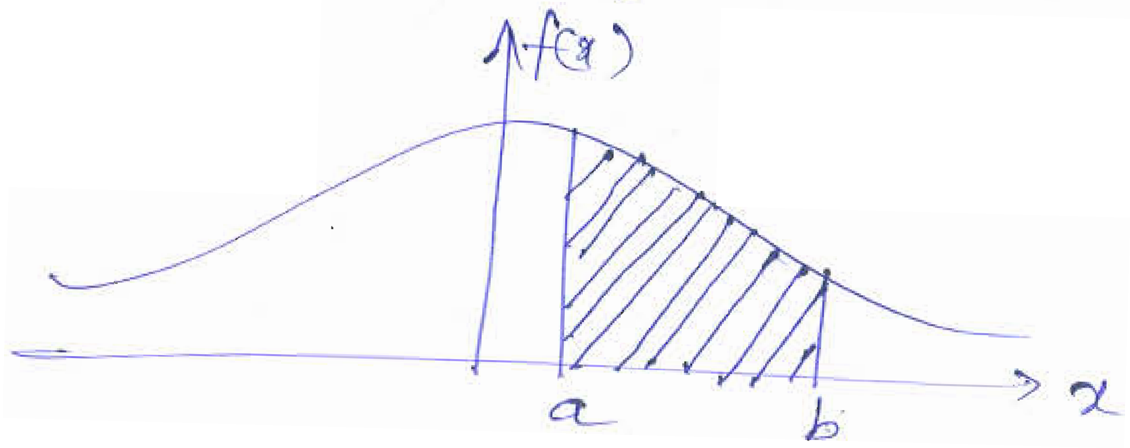
→ $f(x) = 0$ if x is not in R_X.

* Note

$$P(a \le x \le b) = P(a < x \le b) = P(a \le x < b) = P(a < x < b)$$

$$\left[\because \int_a^a f(x)\,dx = 0\right]$$

* Graphical representation of eq (1) is as below -



Cumulative Distribution Function (CDF)

- CDF denoted by $F(x)$ measures the probability that the random variable X assumes a value less than or equal to x.

  ie $F(x) = P(X \le x)$

* If X is discrete,

$$F(x) = \sum_{\text{all } x_i \le x} P(x_i).$$

If X is continuous,

$$F(x) = \int_{-\infty}^{x} f(x)\,dx.$$

* properties of CDF

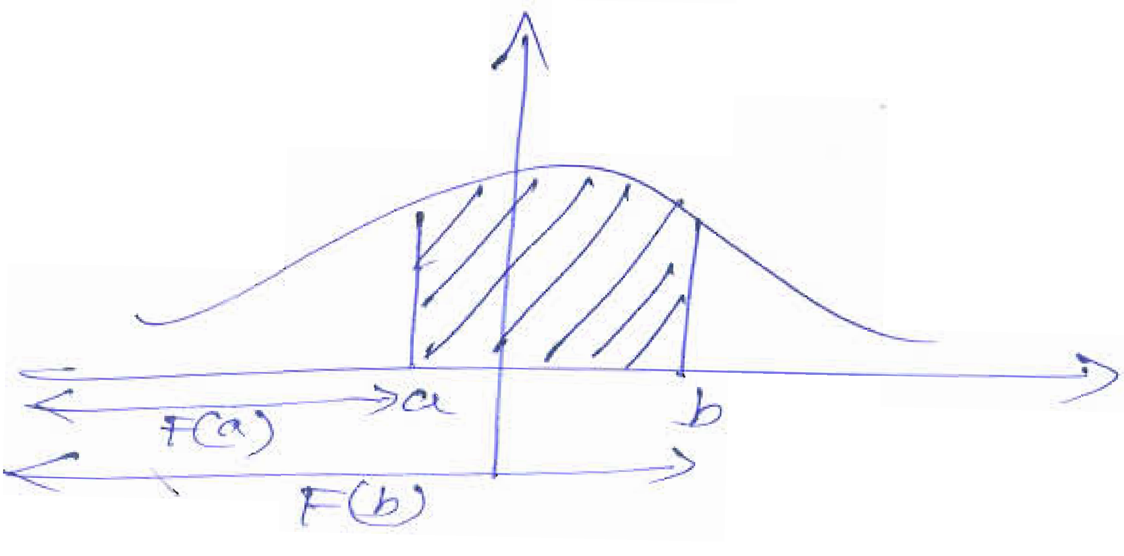
→ $F(x)$ is a non decreasing function.

→ $\lim_{x \to -\infty} F(x) = 0$ [Min value of $F(x)$]

→ $\lim_{x \to \infty} F(x) = 1$ [Max value of $F(x)$]

* Note:

$P(a < X < b) = F(b) - F(a) \quad \forall \ a < b$



Mean [Expectation]

* **Definition**

Mean is the Expected value of the Random variable x.

It is defined as follows:

$$E(x) = \sum_{\text{all } i} x_i \, p(x_i) \longrightarrow \text{if } x \text{ is discrete}$$

$$E(x) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx \longrightarrow \text{if } x \text{ is continuous.}$$

* It is also called as first moment of x.

Generally, n^{th} moment of x can be defined as :-

$$E(x) = \sum_{\text{all } i} x_i^n \cdot p(x_i) \longrightarrow \text{if } x \text{ is discrete}$$

$$E(x) = \int_{-\infty}^{\infty} x^n \cdot f(x) \longrightarrow \text{if } x \text{ is continuous.}$$

Variance

* **Definition.**

Variance of a Random variable x is the deviation from its Expected value (mean).

It is defined as -

$$V(x) = \sigma^2 = E(x^2) - E(x)^2$$

Standard Deviation

$$\sigma = \sqrt{\sigma^2} = \sqrt{V(x)}.$$

Mode

* **Definition**

In discrete case, Mode is the value of x that occurs most frequently.

In continuous case, Mode is the value at which the pdf is maximized.

2. Explain the following Discrete Distributions.

a.) Bernoulli process

b.) Binomial Distributions

c.) Geometric Distributions

d.) Negative Binomial Distributions.

Bernoulli Process.

* **Bernoulli Trial**

Bernoulli trial is an experiment which can be either success or Failure.

Let $X = 1$, if exp. is success
$X = 0$, if exp is failure.

* **Bernoulli process.**

The 'n' Bernoulli trials are called Bernoulli process if -

→ trials are independent
→ Each trial has only two outcomes
→ probability of success remains constant.

*
$$P_j(x_j) = p(x_j) = \begin{cases} p & x_j = 1, \ j = 1, 2 \ldots n \\ 1-p = q & x_j = 0, \ j = 1, 2 \ldots n \\ 0 & \text{otherwise} \end{cases}$$

* Mean, $E(x_j) = p$

* Variance, $V(x_j) = p(1-p)$
$= p \cdot q$

Binomial Distribution.

* The random variable X that denotes the number of success in 'n' Bernoulli trials has a Binomial distribution given by -

$$p(x) = \begin{cases} \binom{n}{x} \cdot p^{x} \cdot q^{n-x}, & x = 0, 1 \ldots n \\ 0, & \text{otherwise} \end{cases}$$

* Mean, $E(X) = np$

* Variance,

$$V(X) = npq$$

Geometric Distributions

* The Random Variable X that denotes the number of trials required to achieve first success in 'n' Bernoulli trials has a Geometric Distribution given by -

$$p(x) = \begin{cases} q^{x-1} \cdot p & x = 0, 1, \ldots n \\ 0 & \text{otherwise} \end{cases}$$

* Mean,

$$E(X) = \frac{1}{p}$$

* Variance

$$V(X) = \frac{q}{p^2}$$

Negative Binomial Distribution

* The Random variable Y that denotes the number of trials required to achieve k^{th} success in 'n' Bernoulli trials has a Negative Binomial Distribution given by -

$$p(y) = \begin{cases} \binom{y-1}{k-1} q^{y-k} \cdot p^k & , \text{ if } y = k, k+1 \ldots\ldots \\ 0 & , \text{ otherwise} \end{cases}$$

* Mean,

$$E(Y) = \frac{k}{p}$$

* Variance,

$$V(Y) = \frac{kq}{p^2}$$

* It is a generalized form of Geometric Distribution.

Example:

Consider 100 Bernoulli trials with success $p = 0.60$ (= 60%) and failure $q = 0.40$ (= 40%)

Find

(a) Probability of no. of success greater than 2.

(b) Probability that total no. of trials required to achieve first success is 10.

(c) Probability that total no. of trials required to achieve 5th success is 20.

(a.) probability of total no. of success greater than 2 is given by

$$p(x>2) = 1 - p(x \le 2)$$

$$= 1 - \sum_{x=0}^{2} \binom{n}{x} p^x \cdot q^{n-x}$$

$$= 1 - \sum_{0}^{2} \binom{100}{x} (0.6)^x \cdot (0.4)^{100-x}$$

$$= 1 - \left[\binom{100}{0} \cdot 1 (0.4)^{100} + \binom{100}{1} (0.6)(0.4)^{99} + \binom{100}{1} \cdot (0.6)^2 \cdot (0.4)^{98} \right]$$

$$= 1 -$$

$$=$$

(b.) probability that total no. of trials required to achieve first success is 10 is given by –

$$P(X=10) = q^{x-1} \cdot P = (0.4)^9 \cdot (0.6)$$

$$P(10) =$$

(c.) Probability that total no. of trials required to achieve 5th success is 20 is given by –

$$P(\cancel{y}=20) = \binom{y-1}{k-1} \cdot q^{y-k} \cdot p^k$$

$$= \binom{19}{4} \cdot (0.4)^{15} \cdot (0.6)^5$$

$$P(20) =$$

3. Explain the following Continuous Distributions

a.) Uniform Distribution

b.) Exponential Distribution.

Uniform Distribution

* A random Variable x is uniformly distributed on the interval (a, b) if its pdf and cdf are given by -

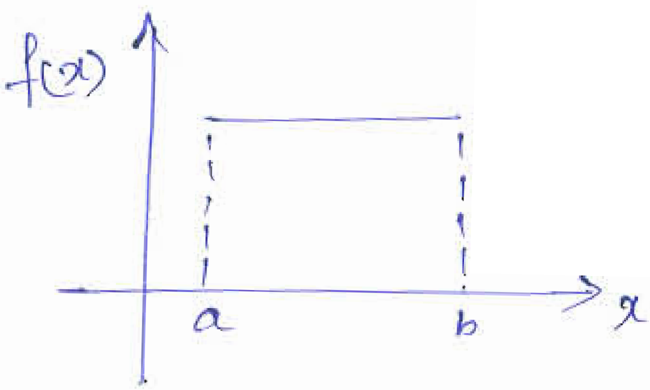
pdf,

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

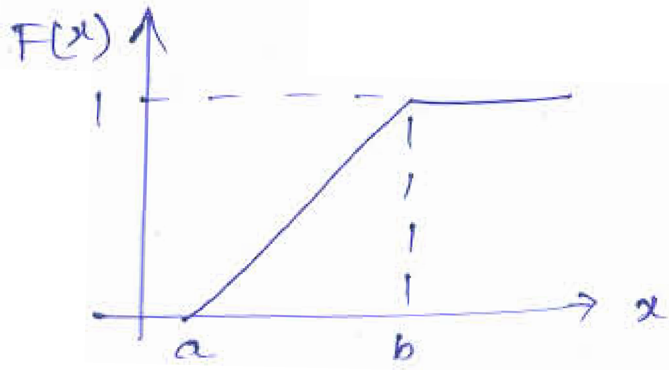
cdf,

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x \geq b \end{cases}$$

* CDF



* PDF



* Note: $P(x_1 < X < x_2) = F(x_2) - F(x_1)$

$$= \frac{x_2 - x_1}{b-a}$$

* Mean, $E(x) = \frac{a+b}{2}$ * Variance, $V(x) = \frac{(b-a)^2}{12}$

- **Example**

If the interval of any Random variable x which is uniformly distributed are :-

$$a = 0$$
$$b = 30.$$

Then,

$$P(0 < x < 15) = F(15) - F(0) = \frac{15-0}{30-0} = \frac{1}{2} = \boxed{0.50}$$

$$P(20 < x < 30) = F(30) - F(20) = \frac{30-20}{30-0} = \frac{1}{3} = \boxed{0.33}$$

Exponential Distribution.

* A random variable x is said to be exponentially distributed with parameter $\lambda > 0$, if its pdf and cdf are given by:-

pdf :-

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

cdf,

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - \lambda e^{-\lambda x} & x \geq 0 \end{cases}$$

* Mean,

$$E(x) = \frac{1}{\lambda}$$

* Variance,

$$V(x) = \frac{1}{\lambda^2}$$

* **Memoryless property.**

Memoryless property of an exponentially distributed random variable states that,
"Future values are not affected by its past values"

In other words,
"Given a full history of the past, we can not determine the future values".

* **Mathematical proof:**

$$P(X > s+t \mid X > s) = \frac{P(X > s+t)}{P(X > s)}$$

$$= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}$$

$$= \frac{e^{-\lambda s} \cdot e^{-\lambda t}}{e^{-\lambda s}}$$

$$\boxed{P(X > s+t \mid X > s) = e^{-\lambda t}}$$

As we can see from this expression, the value of a random variable X does not depends upon the past time 's'. It only depends upon the future time 't'

Example :

If the mean $\lambda = 1/3$ for an exponentially distributed Random Variable x . then

$$P(x > 3) = 1 - P(x \leq 3)$$

$$= 1 - \int_0^3 \lambda e^{-\lambda x} \, dx$$

$$= 1 - [1 - e^{-3/3}]$$

$$= \boxed{0.368}$$

$$P(2 < x < 3) = \int_2^3 \lambda e^{-\lambda x} \, dx$$

$$= \boxed{0.145}$$

$$P(x > 3.5 \mid x > 2.5) = e^{-1/3}$$

$$= \boxed{0.717}$$

4. Explain the poissons process in detail

* Definition:

$N(t)$ is a counting function that represents the number of events occurred in $[0, t]$.

Any counting function $(N(t), t \geq 0)$ is a poissons process with mean rate λ, if -

a.) Events Arrival occurs one at a time

b.) $\{N(t), t>0\}$ has stationary increments.
ie No. of events b/w t and $t+s$ depends only on length of interval s, and not on starting point t.

c.) $\{N(t), t>0\}$ has independent increments.
ie No. of events during non-overlapping time intervals are independent random variables.

* If Arrivals occurs according to a poissons process, then probability that $N(t) = n$ is given by

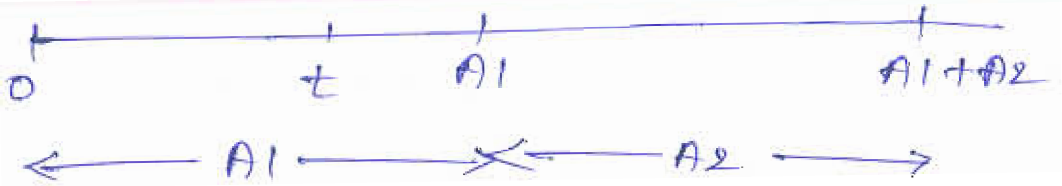
$$p(N(t)=n) = \frac{e^{-\lambda t}.(\lambda t)^n}{n!}$$

* Mean,
$$E(x) = \lambda t = E[N(t)]$$

* Variance
$$V(x) = \lambda t = V[N(t)]$$

Inter Arrival time.

* Consider the inter arrival times of a poissons process $(A_1, A_2 \ldots)$



a.) First Arrival occurs after t, if and only if there are no arrivals in the interval $(0,t)$

Hence,

$$P(A_1 > t) = P(N(t) = 0) = \boxed{e^{-\lambda t}}$$

b.) probability that the first Arrival will occur in $(0,t)$ is given by -

$$P(A_1 \le t) = 1 - P(A_1 > t)$$

$$= \boxed{1 - e^{-\lambda t}}$$

properties of a poissons process

a.) splitting

2.) pooling

splitting

* suppose each event of a poissons process can be classified as type I with probability p, and type II with probability $1-p$, then,

$N(t) = N_1(t) + N_2(t)$.

$$
\begin{tikzcd}
 & N_1(t) \sim poi(\lambda p) \\
N(t) \sim poi(\lambda) \arrow[r, "\lambda"] & \cdot \arrow[u, "\lambda p"] \arrow[d, "\lambda(1-p)"'] \\
 & N_2(t) \sim poi(\lambda(1-p))
\end{tikzcd}
$$

pooling

* It is a opposite situation from random splitting ie. two poissons process are pooled together.

$N_1(t) + N_2(t) = N(t)$

$$
\begin{tikzcd}
N_1(t) \sim poi(\lambda 1) \arrow[dr, "\lambda 1"] & & \\
 & \cdot \arrow[r, "\lambda 1 + \lambda 2"] & N(t) \sim poi(\lambda 1 + \lambda 2) \\
N_2(t) \sim poi(\lambda 2) \arrow[ur, "\lambda 2"'] & &
\end{tikzcd}
$$

Non Stationary Poissons Process

* If any counting function $\{N(t), t>0\}$ does not satisfy the 2^{nd} condition of poissons process, but satisfies the other two, then the counting function $N(t)$ is called Non Stationary poissons process.

* Definition:

Non Stationary poissons process (NSPP) is a poissons process without the stationary increments, characterized by $\lambda(t)$, the arrival rate at time t.

* The expected ~~value~~ no. of arrivals by time t, is given by:-

$$\Lambda(t) = \int_0^t \lambda(s)\,ds.$$

* Example,

let $\lambda(t) = \begin{cases} 2, & 0 < t < 4 \\ 1/2, & 4 < t < 8 \end{cases}$

then $\Lambda(t) = \begin{cases} 2t, & 0 < t < 4 \\ t/2 + 6, & 4 < t < 8 \end{cases}$ $\because \Lambda(t) = \int_0^t \lambda(s)\,ds$

$$P[N(6) - N(3) = k] = P[N(\Lambda(6)) - N(\Lambda(3)) = k]$$
$$= P[N(9) - N(6) = k]$$
$$= \frac{e^{9-6} \cdot (9-6)^k}{k!}$$
$$= \boxed{\frac{e^3 \cdot 3^k}{k!}}$$

ASHOK KUMAR. K
9742024066
celestialcluster@gmail.com.

UNIT 5: RANDOM NUMBER GENERATION
RANDOM VARIATE GENERATION

1. What are pseudo Random Numbers? Describe the properties of Random Numbers.

2. What are the possible errors and considerations in the generation of Random Number routines.

3. Explain the two techniques for generating the Random Numbers.

4. Explain the Auto Correlation test for testing the RN for Independence.

5. What is Inverse transform technique? Derive an expression for exponential Distribution.
   (or) Suggest a step by step procedure to generate random variates using Inverse Transform technique for exponential distribution.

6. Explain Acceptance - Rejectance Technique.

7. Given the uniform distribution on $(1, 2 .. k)$ with pmf $p(x) = 1/k$, $x = 1, 2 .. k$.
   Generate the random variables for five random nos $0.81, 0.12, 0.34, 0.56, \& \ 0.9$.
   use $k = 10$.
   Derive the formula used.

ASHOK KUMAR. K
9742024066
celestialcluster@gmail.com

1. What are pseudo Random Numbers? Describe the properties of Random Numbers.

Pseudo Random Numbers.

* Definition:

pseudo Random Numbers are set of values or elements that is statistically random, but it is derived from a known starting point and is typically repeated over and over.

"pseudo", because generating numbers using a known method removes the potential for true randomness.

* properties of Random Numbers.

a.) Uniformity
b.) Independence

* Additional properties

c.) Maximum Density
d.) Maximum period.

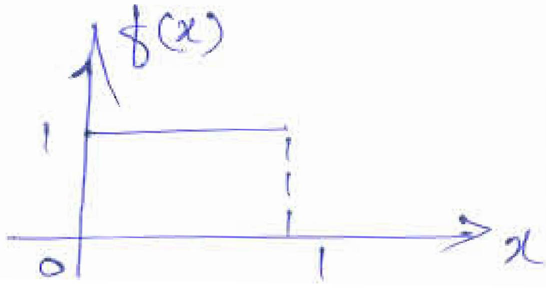
Uniformity → they are equally probable every where

Independence → current value of a Random variable has no relation with previous values.

ie, Each Random Number R_i must be an independent sample drawn from a continuous uniform distribution b/w 0 and 1.

Its pdf is -

$$f(x) = \begin{cases} 1 & 0<x<1 \\ 0 & \text{otherwise} \end{cases}$$



Mean, $E(x) = 1/2$

Variance, $V(x) = 1/12$

<u>Additional properties</u>

→ <u>Maximum Density</u>, meant that the values assumed by R_i, $i = 1, 2 \ldots$ leave no longer gaps on $[0, 1]$

→ <u>Maximum Period</u>.
To help achieve maximum density, the random number generator should have the largest possible period.

2. What are the possible errors and considerations in the generation of Random Number routines?

Possible Errors or problems in the generation of Random Numbers

a.) The generated numbers might not be uniformly distributed.

b.) They may be discrete valued instead of continuous valued.

c.) The mean of the generated numbers may be too high or too low.

d.) The variance of the generated numbers may be too high or too low.

e.) There might be dependency. Like:-

→ Auto Correlation b/w Numbers

→ Numbers successively higher or lower than adjascent numbers

→ Several numbers above the mean followed by several numbers below the mean.

Important considerations in Random Number routines

a.) The routine should be <u>FAST</u>.
Because, the simulation could require many millions of Random numbers.

b.) The routine should be <u>PORTABLE</u> to different computers, and ideally to different programming language.

c.) The routine should have a sufficiently <u>LONG CYCLE</u>.
The cycle length or period represents the length of random number sequence before the previous numbers begin to repeat themselves in earlier order.

d.) The random numbers should be <u>REPLICABLE</u>.
Given the starting point, it should be possible to generate the same set of Random numbers completely independent of system that is being simulated.

e.) Generated Random Numbers should closely approximate the ideal <u>STATISTICAL PROPERTIES</u> of Uniformity and Independence.

5. **Explain the two techniques for generating the Random Numbers.**

There are three techniques for generating the Random Numbers

a.) Linear Congruential Method (LCM)
b.) Combined Linear Congruential Method (CLCM)
c.) Random Number streams.

Linear Congruential Method (LCM)

* LCM produces a sequence of integers $x_1, x_2 \ldots$ between 0 and $m-1$ following a recursive relationship -

$$\boxed{X_{i+1} = (a X_i + c) \bmod m} \qquad i = 0, 1, 2 \ldots$$

where, $X_0 \rightarrow$ initial value, called seed.
$a \rightarrow$ Multiplier
$c \rightarrow$ Increment
$m \rightarrow$ modulus.

If $c \neq 0$, it is called Mixed Congruential method.
If $c == 0$, it is called Multiplicative Congruential method.

* Random Numbers between 0 and 1 can be generated by -

$$\boxed{R_i = \frac{X_i}{m}} \qquad i = 1, 2 \ldots$$

* **Example**,

Let $x_0 = 27$, $a = 17$, $c = 43$, $m = 100$

then, the Random numbers are generated using LCM as follows

$x_1 = (a x_0 + c) \mod m$

$= (27 \cdot 17 + 43) \mod 100$

$\boxed{x_1 = 2} \Rightarrow r_1 = x_1/m = 2/100$, $\boxed{r_1 = 0.02}$

$x_2 = (a \cdot x_1 + c) \mod m$

$= (17 \cdot 2 + 43) \mod 100$

$\boxed{x_2 = 77} \Rightarrow r_2 = x_2/m = 77/100 = \boxed{0.77}$

$x_3 = (a \cdot x_2 + c) \mod m$

$= (17 \cdot 77 + 43) \mod 100$

$\boxed{x_3 = 52} \Rightarrow r_3 = x_3/m = 52/100 = \boxed{0.52}$.

and so, on.

* **NOTE 1:** LCM was proposed by Lehmar in 1951

NOTE 2: selection of values of a, c, m & x_0 affects the properties and cycle length.

NOTE 3:

Rule 1: If $c \neq 0$, $M = 2^b$, $a = 1 + 4k$, c is relatively prime to 'm', then $\boxed{\text{Max period} = p = m = 2^b}$

Rule 2: If $c = 0$, $M = 2^b$, $x_0 = \text{odd}$, $a = 3 + 5k$ or $a = 8 + 5k$, then, $\boxed{\text{Max period} = p = m/4}$

Rule 3: If $c = 0$, $M = \text{prime}$, $a^{k} - 1$ is divisible by m, then, $\boxed{\text{Max period} = p = m - 1}$

Combined Linear Congruential Generator (CLCG)

* **Reason**

Longer period generator is needed because of increasing complexity of simulated systems.

* **Approach**

combine two or more multiplicative congruential generators.

* Let $X_{i,1}, X_{i,2}, \ldots X_{i,k}$ be the i^{th} output from k different multiplicative congruential generators,

Then,

combined generators are of the form -

$$X_i = \left[\sum_{j=1}^{k} (-1)^{j-1} X_{i,j}\right] \bmod (m_1 - 1)$$

with,

$$R_i = \begin{cases} X_i/m_1 & , \quad X_i > 0 \\ \dfrac{m_1 - 1}{m_1} & , \quad X_i = 0 \end{cases}$$

$$\left.\begin{array}{l}\text{Maximum} \\ \text{possible} \\ \text{period}\end{array}\right\} P = \frac{(m_1 - 1)(m_2 - 1) \cdots (m_k - 1)}{2^{k-1}}$$

4. **Explain Auto Correlation test for Testing the Random Numbers for Independence.**

Auto Correlation Test

- Tests for Auto Correlation are concerned ~~about~~ with the dependence b/w numbers in a sequence.
- This test requires the computation of autocorrelation b/w every m numbers ($m \rightarrow$ lag) starting with i^{th} number.

Thus, Autocorrelation ρ_{im} b/w the following numbers would be of interest:

$$R_i,\ R_{i+m},\ R_{i+2m} \cdots R_{i+(M+1)m}$$

The value M is the largest integer such that,

$$i + (M+1)m \le N \quad , \quad N \rightarrow \text{total no. of RNs.}$$

- A non zero Auto correlation implies a lack of independence.

So, the following two-tailed test is appropriate -

$H_0 : \rho_{im} = 0$; if nos are independent

$H_1 : \rho_{im} \neq 0$, if nos are dependent.

- For large values of M, the distribution of the estimator of ρ_{im}, denoted by $\hat{\rho}_{im}$ is approximately normal.
- Test statistics can be formed as -

$$\boxed{Z_0 = \frac{\hat{\rho}_{im}}{\sigma_{\hat{\rho}_{im}}}}$$

5. **What is Inverse transform Technique? Derive an expression for exponential Distribution.**

Inverse Transform Technique

* The Basic principle of Inverse transform technique is to find the inverse function of F, ie F^{-1} such that,

$$\boxed{F \cdot F^{-1} = F^{-1} \cdot F = 1}$$

* F^{-1} denotes the solution of the equation $R = F(x)$ in terms of R, (not $1/F$)

* Ex:- a.) Inverse of $y = x$ is $x = y$.

  b.) Inverse of $y = 2x+1$ is $x = \frac{y-1}{2}$

Expression for Exponential Distribution.

* pdf of Exponential distribution is -

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

* cdf of Exponential distribution is -

$$F(x) = \int_{-\infty}^{x} f(x)\,dx = \begin{cases} 1 - e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

* The idea is to solve $y = 1 - e^{-\lambda x}$ for x.

* <u>Steps involved are as follows</u>

Step 1:
Compute the CDF of the desired random variable X.
CDF of exponential distribution is -

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

Step 2
Set $F(X) = R$ on the range of x.
For Exponential Distribution,

$$R = 1 - e^{-\lambda x} \quad \text{on the range of } x \geq 0.$$

Step 3
Solve the equation $F(X) = R$ for x in terms of R.
In case of Exponential Distribution,

$$1 - e^{-\lambda x} = R$$

$$e^{-\lambda x} = 1 - R$$

$$x = \frac{-1}{\lambda} \ln(1-R)$$

$$\boxed{x = \frac{-1}{\lambda} \ln R}$$ [∵ Both $1-R$ and R are uniformly distributed on $(0,1)$]

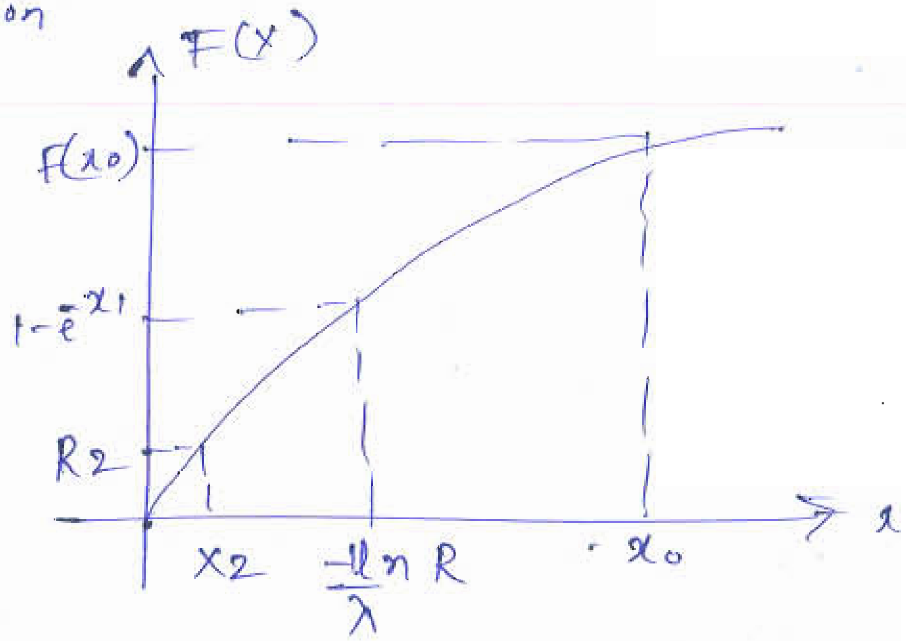
ie $x = F^{-1}(R)$

Step 4:
Generate the uniform Random numbers as needed $R_1, R_2 \ldots$ and compute the Random variates by -

$$X_i = F^{-1}(R_i)$$

$$\boxed{X_i = -\frac{1}{\lambda} \ln R_i} \quad i = 1, 2, 3 \ldots$$

* Fig below shows graphical representation of inverse transform technique applied for exponential distribution



* To generate a value x_1 with CDF $F(x)$, a random no. R_1 b/w 0 and 1 is generated and then a horizontal line is drawn from R_1 to the graph of CDF, & then a vertical line is drawn to x-axis to obtain x_1, the desired result.

6. Explain Acceptance - Rejectance Technique.

Acceptance - Rejection Technique

* It is useful particularly when inverse CDF does not exist in closed form.
* Also called thinning.
* To generate Random variates X, uniformly distributed b/w $1/4$ and 1 ie $X \sim U(1/4, 1)$

step 1: Generate a Random No, $R \sim U(0,1)$

Step 2A: If $R > 1/4$, Accept $x = R$. Goto step 3.

step 2B: If $R \leq 1/4$, Reject $x = R$, Return to step 1

step 3: If another Random variate is needed, Repeat the procedure from step 1.
Else, STOP.

* R itself does not have the desired distribution. But R conditioned on the event $(R \geq 1/4)$ does have the desired distribution.

proof,
Take $1/4 \leq a < b < 1$, Then,

$$P\left[(a \leq R < b) \mid (1/4 \leq R \leq 1)\right] = \frac{P(a < R \leq b)}{P(1/4 \leq R \leq 1)}$$

$$= \frac{b-a}{3/4}$$

which is the correct probability for a uniform distribution on $(1/4, 1)$

* Efficiency:

Depends heavily on the ability to minimize the number of Rejections.

Here the probability of rejection is $P(R < 1/4)$

So, $p(\text{success}) = p = 3/4$

$p(\text{failure}) = 1/p - 1 = \underline{\underline{1/3}}$

Mean no. of random numbers R required to generate one random variate X is one more than no. of rejections

Hence, it is $4/3 - 1 + 1 = 1/3 + 1 = \underline{\underline{1.33}}$

7. Given the uniform Distribution on $(1, 2 \ldots k)$ with pmf $p(x) = 1/k$, $x = 1, 2 \ldots k$.
Generate the random variables for five Random nos 0.81, 0.12, 0.34, 0.56, & 0.9.
Use $k = 10$.
Derive the formula used.

Derivation of the formula.

- Consider the discrete uniform distribution on $(1, 2 \ldots k)$ with pmf, $p(x) = 1/k$, $x = 1, 2 \ldots k$.

- Its CDF is given by -

$$F(x) = \begin{cases} 0, & x < 1 \\ 1/k, & 1 < x < 2 \\ 2/k, & 2 < x < 3 \\ \vdots & \\ k-1/k, & k-1 < x < k \\ 1, & k < x \end{cases}$$

Let $x_i = i$, and

$$r_i = p(1) + p(2) + \ldots + p(x_i)$$

$$\Rightarrow r_i = F(x_i)$$

$$= i/k \quad \text{for } i = 1, 2 \ldots k.$$

WKT,

$$\left[F(x_{i-1}) = r_{i-1}\right] < R < \left[F(x_i) = r_i\right]$$

From this,
If the generated Random number R satisfies

$$r_{i-1} < R < r_i$$

$$\frac{i-1}{K} < R < \frac{i}{K}$$

Multiply by K

$$i-1 < RK < i$$

(or) $RK < i < RK+1$ ——— (2)

Let $\lceil y \rceil$ denote the smallest integer $\geq y$ this notation and eq (2) yield a formula for generating x

ie $\boxed{x = \lceil RK \rceil}$

Given,

$R_1 = 0.81$ $\quad x_1 = \lceil R_1 \cdot K \rceil = 0.81 \times 10 = \lceil 8.1 \rceil = \boxed{9}$

$R_2 = 0.12$ $\quad x_2 = \lceil 1.2 \rceil = \boxed{2}$

$R_3 = 0.34$ $\quad x_3 = \lceil 3.4 \rceil = \boxed{4}$

$R_4 = 0.56$ $\quad x_4 = \lceil 5.6 \rceil = \boxed{6}$

$R_5 = 0.9$ $\quad x_5 = \lceil 9 \rceil = \boxed{9}$

ASHOK KUMAR. K
9742024066
celestialcluster@gmail.com

USN | | | | | | | | | |

2002 SCHEME

CS65

Sixth Semester B.E. Degree Examination, June-July 2009
System Simulation and Modeling

Time: 3 hrs. Max. Marks:100

Note: *Answer any FIVE full questions.*

1 a. What is simulation? State any two of its merits and two limitations. State any two situations where simulation can be used. (05 Marks)
b. Discuss the types of models of a system. (05 Marks)
c. Explain various steps in a simulation study. Write the flow chart for simulation study. (10 Marks)

2 a. Describe a queuing system with respect to arrival and service mechanisms, system capacity, queue discipline, flow diagrams of arrival and service events. (08 Marks)
b. A newspaper seller classifies his days into "good" and "bad" ones with probability 0.4 and 0.6 respectively. The amounts of newspaper sold are given by the distributions below.

	Good	Copies sold	Probability	R-D	Bad	Copies sold	Probability	R-D
1-4	0.05	150	0.05	1-5	0.1	150	0.1	1-10
4-10	0.15	200	0.1	6-15	0.4	200	0.3	11-40
	0.35	250	0.2	16-35	0.8	250	0.4	41-80
	0.70	300	0.35	36-70	0.95	300	0.15	81-95
	1.00	350	0.3	71-100	1.00	350	0.05	96-100

He can buy a copy of the newspaper himself by 1 euro and he sell it with the price of 1.8 euros. Unsold copies must be thrown away. Based on 5 days of simulation calculate the profit of the newspaper seller. Instead of 250 newspapers per day if 300 newspapers per day are purchased will it be more profitable.

Random digits for type of day:	3	0	4	7	8
Random digits for number of copies sold:	70	37	93	07	45

(12 Marks)

3 a. Explain the various steps used in Time-advance algorithm. (04 Marks)
b. Six trucks are used to haul coal from the entrance of a small mine to the railroad. Each truck is loaded by one of two loaders. After loading, a truck immediately moves to the scale, to be weighed. Both the loaders and the scale have FCFS waiting line for trucks. After being weighed, a truck begins a travel time and then after ward returns to the loader queue. It is assumed that 5 of the trucks are at the loaders and one is at the scale at time 0. The activity times are given in the following table:

Loading time (min)	10	5	15	5	10
Weighing time (min)	12	16	12	12	12
Travel time (min)	40	60	80	60	40

Simulate the system for 25 minutes, estimate the loader and scale utilization. (16 Marks)

4 a. Define a random number. Explain statistical properties of random numbers with example. (05 Marks)
b. Generate 4 three digit random numbers using multiplicative congruential method with $X_0=117$, $a = 43$ and $m = 1000$. (06 Marks)
c. The sequence of numbers 0.54, 0.73, 0.98, 0.11, 0.68 has been generated. Use the Kolmogorov-Smirnov test with $\alpha = 0.05$ to determine if the hypothesis that the numbers are uniformly distributed on the internal [0, 1] can be rejected. Compare $F(x)$ and $S_N(X)$ on a graph. $D_{0.05,5} = 0.565$. (09 Marks)

5 a. When to use random variate generation? What is the difference between random number generation and random variate generation? Explain with example. (05 Marks)

b. Explain the inverse transformation technique of producing random variates for exponential distribution. Generate exponential variates X_i with mean 1. Given random numbers R_i=0.1306, 0.0422, 0.6597, 0.7965, 0.7696. (08 Marks)

c. What is acceptance-rejection technique? Generate 3 Poisson variates with mean $\alpha = 0.2$. Use the following random nos. 0.4357, 0.4146, 0.8353, 0.9952, 0.8004. (07 Marks)

6 a. What is the need for input modeling? Explain the steps involved in the development of a useful model for a given set of input data. (06 Marks)

b. The time required for 50 different employees to compute and record the number of hours during the week was measured with the following results in minutes. Use Chi-square test, to test the hypothesis that these service times are exponentially distributed. Take the number of class intervals as $K = 6$, $\alpha = 0.05$.

Employee	Time	Employee	Time	Employee	Time	Employee	Time	Employee	Time
1	1.88	11	3.53	21	1.42	31	0.39	41	0.80
2	1.54	12	0.53	22	1.28	32	0.34	42	5.50
3	1.90	13	1.80	23	0.82	33	0.01	43	4.91
4	0.15	14	0.79	24	2.16	34	0.10	44	0.35
5	0.02	15	0.21	25	0.05	35	1.10	45	0.36
6	2.81	16	0.80	26	0.04	36	0.24	46	0.90
7	1.50	17	0.26	27	1.49	37	0.26	47	1.03
8	0.53	18	0.63	28	0.66	38	0.45	48	1.73
9	2.62	19	0.36	29	2.03	39	0.17	49	0.38
10	2.67	20	2.03	30	1.00	40	4.29	50	0.48

Use $\chi^2_{0.05,4} = 9.49$ (14 Marks)

7 a. Explain in detail about the model building, verifying and validation process through a diagram. (10 Marks)

b. What is output analysis? State its purpose. Explain point estimation and interval estimation. (10 Marks)

8 a. Discuss the concept of high-level computer simulations by sketching a simulation model at a computer system that services requests from the world wide web. (08 Marks)

b. Explain CPU and memory simulation. (06 Marks)

c. Explain about simulation tools. (06 Marks)

* * * * *

SET-4 8

USN

2002 SCHEME

CS65

Sixth Semester B.E. Degree Examination, Dec.08 / Jan.09
System Simulation and Modeling

Time: 3 hrs. Max. Marks:100

Note : *Answer any FIVE full questions.*

1 a. With an example, define a model of a system. Give the classification of different types of models of a system. (04 Marks)

b. With necessary example, state any two situations where simulation is not appropriate tool to use. (04 Marks)

c. With a neat flow-chart, briefly explain the different steps involved in a simulation study. (12 Marks)

2 a. Explain any four characteristics of a queueing system. (08 Marks)

b. A small grocery store has only one checkout counter. Customers arrive at this counter at random from 1 to 10 minutes apart. Each possible value of interarrival time has the same probability of occurrence equal to 0.10. The service times vary from 1 to 6 minutes apart with probabilities shown below.

Service time :	1	2	3	4	5	6
Probability	0.10	0.20	0.30	0.25	0.10	0.05

Develop simulation table for 10 customers and find the following:
i) The average time between arrivals.
ii) The probability that a customer has to wait in the queue.
iii) The average service time.
Random digits for arrivals : 91, 72, 15, 94, 30, 92, 75, 23, 30
Random digits for service times : 84, 10, 74, 53, 17, 79, 91, 67, 89, 38 (12 Marks)

3 a. With respect to discrete event simulation, differentiate between the terms activity and delay. (06 Marks)

b. What are pseudo random numbers? List the errors, which occur during the generation of pseudo random numbers. (06 Marks)

c. Use linear congruential method to generate a sequence of three random numbers for $X_0 = 27$, $a = 8$, $C = 47$ and $m = 100$. (08 Marks)

4 a. Consider the 60 two-digit numbers in the sequence given below. Test whether the 2^{nd}, 9^{th}, 16^{th}, ... numbers in the sequence are autocorrelated, where $\alpha = 0.05$ (10 Marks)

0.30	0.48	0.36	0.01	0.54	0.34	0.96	0.06	0.61	0.85
0.48	0.86	0.14	0.86	0.89	0.37	0.49	0.60	0.04	0.83
0.42	0.83	0.37	0.21	0.90	0.89	0.91	0.79	0.57	0.99
0.95	0.27	0.41	0.81	0.96	0.31	0.09	0.06	0.23	0.77
0.73	0.47	0.13	0.55	0.11	0.75	0.36	0.25	0.23	0.72
0.60	0.84	0.70	0.30	0.26	0.38	0.05	0.19	0.73	0.44

Table Q4 (a)

b. Explain the inverse transformation technique of producing random variates for exponential distribution. (05 Marks)

c. Generate three Poisson variates with mean $\alpha = 0.2$. (05 Marks)

5 a. List the steps involved in the development of a useful model of input data. (04 Marks)

b. Explain how the method of histograms can be used to identify the shape of a distribution. (06 Marks)

c. Explain with a neat diagram, the model building, verification and validation. (10 Marks)

6 a. Explain the types of simulation with respect to output analysis. Give an example. (06 Marks)
b. Discuss in brief the output analysis for steady-state simulations. (06 Marks)
c. Discuss about point and interval estimations. (08 Marks)

7 a. Briefly explain the process-oriented and event-oriented simulation tools. (10 Marks)
b. Explain the concept of CPU simulation and memory simulation. (10 Marks)

8 Write short notes on:
a. Data collection in input modeling.
b. Model calibration.
c. Queueing notations.
d. Secondary properties of random numbers. (20 Marks)

*****

SET-4 CS65

USN | | | | | | | | | |

Sixth Semester B.E. Degree Examination, June / July 08
System Simulation and Modeling

Time: 3 hrs. Max. Marks:100

Note : *1. Answer any FIVE full questions.*
2. Use of statistical tables is permitted.

1 a. Define simulation, simulation model, entities, measures-of-performance and activities. (05 Marks)
b. List three circumstances under which simulation is the appropriate tool and two circumstances under which simulation is not the appropriate tool. (05 Marks)
c. Explain in brief with a neat figure the steps involved in a simulation study. (10 Marks)

2 a. Explain in brief a simple queuing model and represent it using queuing notation. (05 Marks)
b. List and describe in brief the five elements / characteristics of the queuing system. (05 Marks)
c. A grocery store has one checkout counter. Customers arrive at this check out counter at random from 1 to 8 minutes apart and each inter arrival time has the same probability of occurrence. The service times vary from 1 to 6 minutes with probabilities as given below.

Service (minutes)	1	2	3	4	5	6
Probability	0.10	0.20	0.30	0.25	0.10	0.05

Simulate the arrival of 5 customers and calculate i) Average waiting time for a customer ii) Probability that a customer has to wait iii) Probability of a server being idle iv) Average service time and v) Average time between arrivals. Use the following sequence of random numbers.

Random digits for arrival :	913	727	015	948	309	922
Random digits for service time :	84	10	74	53	17	79

Assume that first customer arrives at time 0. Depict the simulation in a tabular form. (10 Marks)

3 a. Briefly define any five concepts used in discrete event simulation. (05 Marks)
b. Identify the concepts in the following example (i.e example of Q 3(c) drawing relevant figure. (05 Marks)
c. Six dump trucks are used to haul coal from the entrance of a mine to railroad. Each truck is loaded by one of two loaders. After loading, a truck immediately moves to the scale, to be weighed as soon as possible. Both the loaders and the scale have a first – come, first served waiting line for trucks. Travel time from a loader to scale is considered negligible. After being weighed a truck begins travel time (during which time truck unloads) and then afterwards returns to the loader queue. The activities of loading time, weighing time and travel time are given in the following table.

Loading Time	10	5	5	10	15	10	10
Weighing Time	12	12	12	16	12	16	
Travel Time	60	100	40	40	80		

End of simulation is completion of two weighings from the scale. Depict the simulation table and estimate the loader and scale utilizations. Assume that five of the trucks are at the loaders and one is at the scale at time O. (10 Marks)

4 a. Differentiate between truly random numbers and pseudo random numbers. Mention four properties that random numbers should possess. (05 Marks)

b. Using multiplicative congruential method for generating random numbers, list the random numbers and find the period of generator for $a = 13$, $m = 64$ and $X_0 = 2$. (05 Marks)

c. A sequence of 1000 (one thousand) four digit numbers has been generated and an analysis indicates the following combinations and frequencies.
Four different digits = 565; One pair = 392, Two pairs = 17, Three like digits = 24 and remaining are four like digits. Based on the poker test, test whether these numbers are independent. Use level of significance = 0.05. (10 Marks)

5 a. Elaborate the need for generating random variats. Given probability mass function pmf of random variates and a set of uniform random numbers over the range (0,1), describe the method to generate random variates. (10 Marks)

b. Given the uniform distribution on {1, 2, .., k} with pmf $p(x) = \frac{1}{k}$, $x = 1, 2, .., k$, generate the random variates for the five random numbers (0.81, 0.12, 0.34, 0.56 and 0.93). Derive the formula used. Use K = 10 for generating random variates. (10 Marks)

6 a. Explain the need for input modeling and histogram method of identifying the input distribution. (05 Marks)

b. The number of vehicles arriving at a junction in a five minute period was observed for 100 days. The resulting data is as follows:

No. of arrivals	0	1	2	3	4	5	6	7	8	9	10	11
Frequency	12	10	19	17	10	8	7	5	5	3	3	1

It is presumed that the arrivals follow a Poisson distribution with parameter $\alpha = 3.64$. Using Chi –square test, determine whether the assumption that arrivals follow Poisson distribution can be accepted at a 0.05 level of significance.
(Note : Expected values used should be ≥ 5 for calculation and put the values and calculated values in a tabular form). (15 Marks)

7 a. Differentiate between verification and validation of a simulation model. With a neat diagram, explain the relation between model building, verification and validation. (10 Marks)

b. Describe the three step approach which has been used as an aid in the validation process. (10 Marks)

8 Write short notes on:
a. Terminating and steady state simulations (05 Marks)
b. Point estimation of performance parameters. (05 Marks)
c. CPU simulation (05 Marks)
d. Memory simulation. (05 Marks)

*****

SET-4

CS65

USN

6

Sixth Semester B.E. Degree Examination, Dec. 07 / Jan. 08
System Simulation and Modelling

Time: 3 hrs. Max. Marks:100

Note : *1. Answer any FIVE full questions.*
2. Random number Table and statistical Table book may be supplied.

1 a. What is Simulation? State any four merits and demerits of simulation. (10 Marks)
b. Differentiate the following with examples : i) Static and Dynamic model ii) Discrete and continuous system iii) Deterministic and Stochastic model. (06 Marks)
c. State any two situations where simulation can be used with justification. (04 Marks)

2 a. Briefly explain the simulation of Inventory system and the various measures used to evaluate the system. (08 Marks)
b. Prepare a simulation table for a single channel queuing system using event scheduling/time advance algorithm, until the clock reaches time 21, using the inter arrival timer and service timer given below in the order shown. The stopping event will be at time 30.

Inter arrival time (mins)	8	6	1	8	3	8
Service time (mins)	4	1	4	3	2	4

Compute the cumulative statistics for the following :
i) Busy time of server
ii) Maximum Que length.
iii) Total number of customers who spend 4 or more minutes at the counter.
iv) Total number of departures upto the current simulation time. (12 Marks)

3 a. Define any four concepts in discrete event simulation with suitable examples. (04 Marks)
b. Use the multiplicative congiential method to generate a sequence of four three digit random integers for $X_0 = 117$, $a = 43$ and $m = 1000$. (06 Marks)
c. The sequence of numbers 0.54, 0.73, 0.98, 0.11 and 0.68 has been generated. Use the Kolmogorav - Smirnov test with $\alpha = 0.05$ to determine if the hypothesis that its numbers are uniformly distributed on the interval (0,1) can be rejected. (10 Marks)

4 a. Lead times have been found to be exponentially distributed with mean 3.7 days. Generate five random lead time variates from this distribution using Inverse transform technique. Take $R_1 = 0.01$, $R_2 = 0.13$, $R_3 = 0.35$, $R_4 = 0.65$, $R_5 = 0.53$. (10 Marks)
b. Consider discrete distribution with DMF given by $P(x) = \frac{2x}{k(k+1)}$, $x = 1, 2, \ldots k$. Find an expression for finding the values of Random variates 'X' corresponding to Random number 'R'. (10 Marks)

5 a. Explain in detail the four important steps of development of useful 'Input Model'. (10 Marks)

b. The number of vehicles arriving at an intersection in a 5 minute period between 7:00 AM and 7:05 AM was monitored for 5 working days over a 20 week period. The Table below gives the data.

Arrivals per period X_i	0	1	2	3	4	5	6	7	8	9	10	11
Frequency in number of days	12	10	19	17	10	8	7	5	5	3	3	1

i) Construct frequency table and find mean.
ii) Assume Poisson distribution and estimate the parameter 'α'.
iii) Check for Goodness of fit using Chi^2 – test for significance level of 5%. (10 Marks)

6 a. Explain three step approach for validation process as formulated by Nayler and Finger. (12 Marks)
b. Explain Initialization Bias in output analysis of steady state simulation. (08Marks)

7 a. Briefly explain the sequence of pipeline stages in ILP – CPU simulation of computer systems. (10 Marks)
b. Explain LRU stack evolution technique in simulation of computer memory. (10 Marks)

8 Write short notes on :
a. RUNS test.
b. Acceptance -- Rejection technique.
c. Point Estimation.
d. Calibration process in model building. (20 Marks)

*****

CS65

USN | | | | | | | | | | |

SET-4

5

NEW SCHEME

Sixth Semester B.E. Degree Examination, July 2007
Computer Science and Engineering
System Simulation and Modeling

Time: 3 hrs.] [Max. Marks:100

Note : *1. Answer any FIVE full questions.*
2. The statistical tables A.5 and A.6 of "Discrete Event System Simulation" – Jerry Banks book can be provided.

1 a. What is system and system environment? Explain the components of a system with examples. (10 Marks)
b. What are the advantages of simulation? (05 Marks)
c. Discuss the types of models of a system. (05 Marks)

2 a. Explain the calling population, service time and service mechanisms of a queuing system. (08 Marks)
b. A baker bakes 30 dozens of bread loafs each day. Probability distribution of customers is in table1. Customers order 1, 2, 3 or 4 dozens of bread loafs according to the distributions given in table 2. Assume that on each day all the customers order same dozens of bread loafs. The selling price is Rs.5.4/dozen and making cost is Rs.3.8/dozen. The left over bread loafs will be sold for half price at the end of day. Based on 5 days simulation, calculate the profit of the baker. Instead of 30 dozens, if 40 dozens are baked per day will it be more profitable?

Table 1: Probability distribution of customers.

Number of customers/day	8	10	12	14
Probability	0.35	0.30	0.25	0.10

Table 2: Probability distribution of dozens ordered

Number of dozens/ customers	1	2	3	4
Probability	0.40	0.30	0.20	0.10

Random digits for customers -	50	61	73	24	96
Random digits for dozens -	5	3	7	0	8

(12 Marks)

3 a. What are the two categories of activities? Explain the three phases of activity scanning approach. (06 Marks)
b. Prepare a table using event scheduling time advance algorithm for a check out counter. Stop the simulation when fifth customer departs. Estimate mean response time and proportion of customers who spent 4 or more minutes in the system. Event notice must have event type, time and customer number. (14 Marks)

Inter arrival times	8	6	1	8	3	8	....
Service times	4	1	4	3	2	4	....

Contd.... 2

4 a. What is the role of maximum density and maximum period in generation of random numbers? With given seed 45, constant multiplier 21, increment 49 and modulus 40, generate a sequence of five random numbers. (10 Marks)

b. For the following sequence can the hypothesis that the numbers are independent can be rejected on the basis of length of runs up and down when $\alpha = 0.05$.

8

0.34	0.90	0.25	0.89	0.87	0.44	0.12	0.21
0.46	0.67	0.83	0.76	0.79	0.64	0.70	0.81
0.94	0.74	0.22	0.74	0.96	0.99	0.77	0.67
0.56	0.41	0.52	0.73	0.99	0.02	0.47	0.30
0.17	0.82	0.56	0.05	0.45	0.31	0.78	0.05
0.79	0.71	0.23	0.19	0.82	0.93	0.65	0.37
0.39	0.42						

48

chi square

50. N = 50

(10 Marks)

5 a. A sequence of 1000 four digit numbers has been generated and analysis indicates the following combinations and frequencies:

Combination i	Observed frequency O_i
Four different digits	565
One pair	392
Two pairs	17
Three like digits	24
Four like digits	2

Based on Poker test check whether the numbers are independent. Use $\alpha = 0.05$. (10 Marks)

b. Explain inverse transform technique for exponential distribution. Show the corresponding graphical interpretation. (10 Marks)

6 a. Explain the acceptance – rejection technique. Generate 5 Poisson's variates with mean $\alpha = 0.25$. (10 Marks)

b. Explain Chisquare goodness of fit test. Apply it to Poisson assumption with $\alpha = 3.64$, Data size = 100 and

Observed frequency O_i 12 10 19 17 10 8 7 5 5 3 3 1. (10 Marks)

a. With examples explain output analysis. (10 Marks)

b. Explain with a neat diagram model building, verification and validation. (10 Marks)

Write short notes on:

a. Memory simulation
b. High level computer system simulation
c. Point estimation
d. Errors while generating pseudo random numbers. (20 Marks)

*****

CS65

USN | | | | | | SET-4 | |

4

NEW SCHEME

Sixth Semester B.E. Degree Examination, Dec.06 / Jan.07

Computer Science & Engineering

System Simulation & Modeling

Time: 3 hrs.] [Max. Marks:100

Note: *1. Answer any FIVE full questions.*
2. Use of statistical tables is permitted.

1 a. Define simulation. When simulation is the appropriate tool? (07 Marks)
b. Differentiate between the following terms :
i) System and system environment.
ii) Endogenous and exogenous activity.
iii) Event and activity. (06 Marks)
c. In brief, explain the steps in simulation study. (07 Marks)

2 a. Explain simulation of queueing system. (08 Marks)
b. Dr. XYZ is a dentist who schedules all 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 table shows the various categories of work, their probabilities and time actually needed to complete the work.

Categories	Filling	Crown	Cleaning	Extraction	Checkup
Time required (min)	45	60	15	45	15
Probability of category	0.40	0.15	0.15	0.10	0.20

Simulate the dentist's clinic for 3 hours and determine the average waiting of time for patients and total idle time for doctor. Assume that patients show up at the clinic at exactly their scheduled arrival time starting at 8.00 a.m. Use following random numbers for handling the above problem.
40 82 11 34 25 66 (12 Marks)

3 a. Write discrete-event model for single channel queue. (06 Marks)
b. Prepare simulation table using event scheduling algorithm for the arrival of 8 customers using inter-arrival and service times given below, (14 Marks)

Interval times	4	7	8	1	4	2	5	3	1	4
Service times	5	1	2	3	4	9	5	8	6	1

4 a. Define a random number. Explain statistical properties of random numbers. (03 Marks)
b. Using multiplicative congruential method generate 6 random numbers. (03 Marks)
c. Test the following sequence of numbers for uniformity by using chi-square test and independence by using poker test. ($\alpha = 0.05$) (14 Marks)

0.594	0.928	0.515	0.055	0.507	0.351	0.262	0.797	0.788	0.442
0.097	0.798	0.227	0.127	0.474	0.825	0.007	0.182	0.929	0.852

Contd...2

5. a. Derive an expression for random variate for exponential distribution. (05 Marks)
   b. Lead times have been found to be exponentially distributed with mean 3.7 days. Generate 5 random lead times from this distribution. (03 Marks)
   c. How do you identify the distribution with data? Explain histogram method. (06 Marks)
   d. Explain Kolmogorov-Smirnov goodness of fit test. (06 Marks)

6. a. Differentiate between verification and validation of simulation models. (04 Marks)
   b. Briefly explain 3-step approach that aids in the validation process. (06 Marks)
   c. Explain output analysis for terminating simulation. (10 Marks)

7. a. Explain about simulation tools. (10 Marks)
   b. Explain CPU simulation. (10 Marks)

8. Write short notes on,
   a. Simulation of inventory system.
   b. Time advance algorithm.
   c. Runs tests.
   d. Acceptance – rejection technique. (20 Marks)

*****

NEW SCHEME

CS65

2

Reg. No. SET-4

Sixth Semester B.E. Degree Examination, January/February 2006

Computer Science and Engineering

System Simulation and Modelling

Time: 3 hrs.) (Max.Marks : 100

Note: *1. Answer any FIVE full questions.*
2. Tables A.8 of K.S. critical values and A.6 A.3 of cumulative normal distribution are to be supplied from the discrete - event simulation books.

1. (a) Explain the different steps involved in a simulation study. (8 Marks)

(b) Describe a queuing system with respect to arrival and service mechanisms, system capacity, queue discipline, flow diagrams of arrival and service events. (12 Marks)

2. (a) Differentiate between discrete and continuous system. (5 Marks)

(b) A small shop has one check out counter. Customers arrive at this counter at random from 1 to 10 minutes apart. Each possible value of inter arrival time has the same probability of occurrence equal to 0.10. The service times vary from 1 to 6 minutes with probability shown below.

Service time	1	2	3	4	5	6
Probability	0.05	0.10	0.20	0.30	0.25	0.10

Develop simulation table for 10 customers. Find i) average waiting time ii) average service time iii) average time customer spends in the system. Take random digits for arrivals as
91, 72, 15, 94, 30, 92, 75, 23, 30 and for service times are
84, 10, 74, 53, 17, 79, 91, 67, 89, 38 sequentially. (15 Marks)

3. (a) What are the major concepts in discrete - event simulation? (5 Marks)

(b) One company uses 6 trucks to haul manganese ore from Hospet to its industry. There are two loaders, to load each truck. After loading, a truck moves to the weighing scale to be weighed. The queue discipline is FIFO. When it is weighed a truck travels to the industry and returns to the loader queue. The distributions of loading time, weighing time and trovel time ore as fallows :

Loading times	: 10	5	5	10	15	10	10
Weigh times	: 12	12	12	16	12	16	
Travel times	: 60	100	40	40	80		

Calculate the total busy time of both loaders, of the scale average loader and scale utilization. Assume 5 trucks are at the loaders and one is at the scale at time '0' Stopping time $T_E = 64min.$ (15 Marks)

4. (a) Mention the important considerations for the selection of routines to generate random numbers. (5 Marks)

(b) Explain combined linear congruential generators for developing random numbers. (5 Marks)

Contd.... 2

(c) Use the linear congruential method to generate a sequence of four two-digit random numbers, with $X_0 = 27$, $a = 17$, $C = 43$ and $m = 100$! What is the effect of FIFTH two digit random integer on the above numbers? (10 Marks)

5. (a) The sequence of numbers 0.54, 0.73, 0.98, 0.11, 0.68 has been generated. Use the Kolmogonov - Smirnov test with $\alpha = 0.05$ to determine if the hypothesis that the numbers are uniformly distributed on the interval $[0, 1]$ can be rejected. Compare $F(X)$ and $S_N(X)$ on a graph. (8 Marks)

(b) The 50 two digit values are given below.

0.34	0.90	0.25	0.89	0.87	0.44	0.12	0.21	0.46	0.67
0.83	0.76	0.79	0.64	0.70	0.81	0.94	0.74	0.22	0.74
0.96	0.99	0.77	0.67	0.56	0.41	0.52	0.73	0.99	0.02
0.47	0.30	0.17	0.82	0.56	0.05	0.45	0.31	0.78	0.05
0.79	0.71	0.23	0.19	0.82	0.93	0.65	0.37	0.39	0.42

Can the hypothesis that the numbers are independent be rejected on the basis of the length of runs above and below the mean. Take $\alpha = 0.05$ (12 Marks)

6. (a) Explain the inverse transformation technique of producing random variates for exponented distribution. (5 Marks)

(b) Generate 5 poison variates with mean $\alpha = 0.25$. (5 Marks)

(c) Generate 10 three digit random integers using multiplicative congruential method with $X_0 = 117$, $a = 43$, $m = 1000$. (10 Marks)

7. (a) Explain the Chi-square goodness of fit test to reject or accept a candidate distribution. (5 Marks)

(b) Explain the types of simulation with respect to output analysis. Give at least two examples. (10 Marks)

(c) Write a note on model building, verification and validation. (5 Marks)

8. (a) The time required for 50 different employees to compute and record the number of hours worked during the week was measured with the following results in minutes. Use Chi-square test, to test the hypothesis that these service times are exponentially distributed. Take the number of class intervals as $k = 6$, $\alpha = 0.05$.

Contd.... 3

Employee	Time	Employee	Time	Employee	Time
1	1.88	17	0.26	35	1.10
2	0.54	18	0.63	36	0.24
3	1.90	19	0.36	37	0.26
4	0.15	20	2.03	38	0.45
5	0.02	21	1.42	39	0.17
6	2.81	22	1.28	40	4.29
7	1.50	23	0.82	41	0.80
8	0.53	24	2.16	42	5.50
9	2.62	25	0.05	43	4.91
10	2.67	26	0.04	44	0.35
11	3.53	27	1.49	45	0.36
12	0.53	28	0.66	46	0.90
13	1.80	29	2.03	47	1.03
14	0.79	30	1.00	48	1.73
15	0.21	31	0.39	49	0.38
16	0.80	32	0.34	50	0.48
		33	0.01		
		34	0.10		

(10 Marks)

(b) Write short note on any TWO of the following :

i) Process oriented simulation tools

ii) Concept of CPU simulation

iii) Three step process used in validation process

iv) Memory simulation **(2×5=10 Marks)**

** * **

CS 65

USN SET-4

3

NEW SCHEME

Sixth Semester B.E. Degree Examination, July 2006

CS

System Simulation and Modeling

Time: 3 hrs.] [Max. Marks:100

Note: 1. *Answer any FIVE full questions.*
2. *Random Number Table and Statistical Table book may be supplied.*

1 a. Explain the concept of system with any one live example. (05 Marks)
b. Discuss the various ways of modeling of a system. (05 Marks)
c. Briefly discuss about the various steps in a simulation study through an example. (10 Marks)

2 a. Discuss in detail about the various elements of any general queuing system. Further explain the need for simulation in this environment and the various measures used to evaluate the system. (08 Marks)
b. A news paper seller buys news papers for Rs 3.3 each and sells them for Rs 5 each. Papers not sold at the end of the day are sold as scrap for Rs. 0.5 each. Papers can be purchased in bundles of only 10. There are three types of news days viz.: "Good", " Fair" and "poor" with probabilities 0.35, 0.45 and 0.20 respectively. Determine the optimal number of papers by simulating demands for 20 days. (12 Marks)

3 a. Explain in detail the various steps used in Time – Advance algorithm in a discrete event simulation using a live example. (06 Marks)
b. Six trucks are used to haul coal from a mine to the rail road. There are two loaders and one weighing scale. After loading, a truck immediately moves to the scale for weighing and servicing is as per FIFS. After weighing a truck, begins a travel time and then afterwards return to the loader queue with the distribution of travel time as:

Travel Time in minutes (mts)	40	60	80	100
Probability	0.4	0.3	0.2	0.1

Further the distribution of loading time and weighing time are as :

Loading Time in mts.	5	10	15
Probability	0.3	05	0.2

Weighing Time in mts.	12	16
Probability	0.7	0.3

Simulate the system to estimate the loader and scale utilization. (14 Marks)

4 a. Discuss in brief the various problems or errors which occur while generating Pseudo random numbers. (10 Marks)
b. Explain the Two "Goodness of Fit" tests by using an appropriate example. (10 Marks)

5 a. Explain how and what for the inverse transform technique is used to sample from two discrete distributions. (10 Marks)

Contd...2

b. Discuss how the sample mean is estimated under Normal and POISSON distributions. (10 Marks)

6 a. Explain in detail about the model building , verifying and validation in the model building process through a diagram. (08 Marks)

b. The demand and lead time for product 'X' are as follows :

Demand	83	103	96	92	109	106	104	112	97	116
Lead time	4.3	6.5	6.3	4.5	7.3	5.8	6.9	6.9	6.0	6.9

Test whether lead time and demand are dependent or not . Comment. (12 Marks)

7 a. Discuss in brief the output analysis for steady – state simulations. (08 Marks)

b. Explain the C++ code for generating MPP trace. (12 Marks)

8 a. Discuss about point estimation and interval estimation. (08 Marks)

b. Explain in detail the changes in the computerized representation of the system under I-O transformation. (12 Marks)

*****

NEW SCHEME

CS65

1

USN SET-4

Sixth Semester B.E. Degree Examination, July/August 2005

Computer Science Engineering

System Simulation and Modeling

Time: 3 hrs.] [Max.Marks : 100

Note: 1. Answer any FIVE full questions.
2. Statistical tables may be supplied.

1. (a) With an aid of flow diagram, explain various steps in a simulation study. (10 Marks)

(b) Demand for midgets follows the following probability distribution : (10 Marks)

Daily demand	0	1	2	3	4
Probability	0.33	0.25	0.20	0.12	0.10

Stock is examined every 7 days (the plant is in operation every day) and if the stock level has reached 6 units, or less, an order for 10 midgets is placed. The lead time (days until delivery) is probabilistic and follows the following distribution

Lead time (days)	1	2	3
Probability	0.3	0.5	0.2

When the simulation begins, it is the beginning of the week, 12 midgets are on hand, and no orders have been backordered. (Back ordering is allowed). Simulate 6 weeks of operation of this system. Determine different parameters to analyse the system.

2. (a) With an illustrative example, explain the simulation of queuing system. (10 Marks)

(b) Prepare a table using event scheduling /time advance algorithm, until the clock reaches time 15, using the interarrival and service times given below in the order shown. The stopping event will be at time 30. (10 Marks)

Interval times	4	5	2	8	3	7
Service times	5	3	4	6	2	7

3. (a) Explain he linear congruential method for generating random numbers and generate three random members using above method with $X_0 = 27$, $a = 17$, $c = 43$ and $m = 100$. (10 Marks)

(b) Briefly explain the various tests used for testing the random numbers for their desirable properties. (10 Marks)

4. (a) What is acceptance-Rejection technique? Generate three Poisson variates with mean $\alpha = 0.2$. (10 Marks)

(b) Enlist the steps involved in development of a useful model of input data. (4 Marks)

(c) Describe how the method of histograms can be used to identify the shape of a distribution. (6 Marks)

Contd.... 2

5. (a) Explain the chi-square goodness of fit test to accept or reject a candidate distribution. (10 Marks)

(b) Briefly explain the three-step approach, that aids in the validation process. (10 Marks)

6. (a) Discuss how the performance of a simulated system is measured and estimated, with suitable illustrations. (10 Marks)

(b) With illustrative examples, describe the output analysis for steady state simulations. (10 Marks)

7. (a) Briefly explain the process oriented and event oriented simulation tools. (10 Marks)

(b) Discuss the concepts of high-level computer simulations by sketching a simulation model at a computer system that services requests from the WWW (World wide Web). (10 Marks)

8. Write short notes on the following :

(a) Advantages and disadvantages of simulation

(b) World views

(c) Model building

(d) Memory simulation. (4 × 5=20 Marks)

** * **

USN | | | | | | | | | | |

06CS82

2

Eighth Semester B.E. Degree Examination, December 2010

System Modeling and Simulation

Time: 3 hrs. Max. Marks:100

Note: *Answer any FIVE full questions, choosing atleast TWO questions from each part.*

PART – A

1 a. When is a simulation an appropriate tool? When is it not? (12 Marks)
b. Explain the various components of simulation with an example. (08 Marks)

2 a. Explain the model of 'single channel queue' in detail. (12 Marks)
b. What is list processing? Explain the basic operations of list processing. (08 Marks)

3 a. Briefly explain the various probability terminologies and concepts. (12 Marks)
b. What is Poisson process? Mention the properties of Poisson process. (08 Marks)

4 a. Explain the various steady state parameters of M/G/1 queue. (08 Marks)
b. Explain the service times and server mechanics used in queuing system with an example. (08 Marks)
c. What is networks of queue? Mention the general assumptions for a stable system with infinite calling population. (04 Marks)

PART – B

5 a. Briefly explain the various techniques used to generate random numbers. (12 Marks)
b. Explain any two inverse transform techniques. (08 Marks)

6 a. Mention the important points to be noted while collecting data. (08 Marks)
b. Briefly explain the suggested estimators for distributions often used in simulation. (12 Marks)

7 a. Briefly explain the confidence – interval estimation method. (10 Marks)
b. Explain the two methods to specify the initial conditions in steady state simulation. (10 Marks)

8 a. Differentiate the processes of verification and validation. (04 Marks)
b. Explain the 3 steps involved in model building. (06 Marks)
c. Explain the iterative process of calibrating a model. (10 Marks)

*****

Important Note : 1. On completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages.
2. Any revealing of identification, appeal to evaluator and /or equations written eg, 42+8 = 50, will be treated as malpractice.

USN | | | | | | | | | |

06CS82

Eighth Semester B.E. Degree Examination, May/June 2010
System Modeling and Simulation

Time: 3 hrs. Max. Marks:100

Note: *Answer any FIVE full questions, selecting at least TWO questions from each part.*

Important Note : 1. On completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages.
2. Any revealing of identification, appeal to evaluator and /or equations written eg, 42+8 = 50, will be treated as malpractice.

PART- A

1 a. What is simulation? Explain with flowchart, the steps involved in simulation study. (10 Marks)

b. Differentiate between continuous and discrete systems. (05 Marks)

c. What is system and system environment? List the components of a system, with example. (05 Marks)

2 a. A grocery store has one checkout counter. Customers arrive at this checkout counter at random from 1 to 8 minutes apart and each interval time has the same probability of occurrence. The service times vary from 1 to 6 minutes, with probability given below :

Service (minutes)	1	2	3	4	5	6
Probability	0.10	0.20	0.30	0.25	0.10	0.05

Simulate the arrival of 6 customers and calculate :
Average waiting time for a customer
Probability that a customer has to wait
Probability of a server being idle
Average service time and
Average time between arrival.
Use the following sequence of random numbers :

Random digit for arrival	913	727	015	948	309	922
Random digit for service time	84	10	74	53	17	79

Assume that the first customer arrives at time 0. Depict the simulation in a tabular form. (10 Marks)

b. Briefly define any four concepts used in discrete event simulation. (04 Marks)

c. Explain event scheduling algorithm by generating system snapshots at clock=t and clock = t_1. (06 Marks)

3 a. Six dump trucks are used to have coal from the entrance of a mine to a railroad. Each truck is loaded by one of the two loaders. After loading, a truck immediately moves to the scale, to be weighed as soon as possible. Both the loader and the scale have first-come first-served waiting line for trucks. Travel time from a loader to scale is considered negligible. After being weighed, a truck begins travel time [during which time truck unloads] and then afterwards return to loader queue. The activities of loading, weighing and travel time are given in the following table :

Loading time	10	5	5	10	15	10	10
Weighing time	12	12	12	16	12	16	
Travel time	60	100	40	40	80		

Q-11

End of simulation is completion of two weighings from the scale. Depict the simulation table and estimate the loader and scale utilizations. Assume that five of the trucks are at the loaders and one is at the scale at time 0. (05 Marks)

b. Define a discrete random variable. Explain the binomial distribution. (05 Marks)

c. A production process manufactures alternators for outboard engines used in recreational boating. On the average, 1% of the alternators will not perform up to the required standards when tested at the engine assembly plant. When shipment of 100 alternators is received at the plant, they are tested, and if more than two are non confirming; the shipment is returned to the alternators manufacturer. What is the probability of returning a shipment? (10 Marks)

4 a. Explain the characteristics of a queuing system. List different queuing notations. (10 Marks)

b. A tool crib has exponential interarrival and service times, and it serves a very large group of mechanics. The mean time between arrivals is 4 minutes. It takes 3 minutes on the average for a tool crib attendant to service a mechanic. The attendant is paid \$ 10 per hour and the mechanic is paid \$ 15 per hour. Would it be advisable to have a second tool-crib attendant? (10 Marks)

PART- B

5 a. What are pseudo random numbers? What are the problems that occur while generating pseudo random numbers? (06 Marks)

b. Explain combined linear congruential method for random number generation. (06 Marks)

c. The sequence of numbers 0.54, 0.73, 0.98. 0.11 and 0.68 has been generated. Use the Kolmogorov-Smirnov test with $\alpha = 0.05$ to determine if the hypothesis that the numbers are uniformly distributed on the interval [0, 1] can be rejected. (08 Marks)

6 a. Suggest a step by step procedure to generate random variates using inverse transform technique for exponential distribution. (06 Marks)

b. Enlist the steps involved in development of a useful model of input data. (04 Marks)

c. Records pertaining to the monthly number of job-related injuries at an underground coal mine, were being studied by a federal agency. The values for the past 100 months were as follows :

Injuries per month	0	1	2	3	4	5	6
Frequency of occurrence	35	40	13	6	4	1	1

i) Apply the chi-square test to these data to test the hypothesis, that, underlaying distribution is Poisson. Use a level of significance of $\alpha = 0.05$.

ii) Apply the chi-square test to these data to test the hypothesis, that, the distribution is Poisson with mean 1.0. Again let $\alpha = 0.05$. (10 Marks)

7 a. Briefly explain the measure of performance of a simulation system. (10 Marks)

b. Explain the distinction between terminating or transient simulation and steady state simulation. Give examples. (10 Marks)

8 a. Explain with a neat diagram, model building, verification and validation process. (10 Marks)

b. Describe the three steps approach to validation by Naylor and Finger. (10 Marks)

* * * * *