PAGE NO.:
DATE / /

Chapter-7

• Estimation Method

→ Estimates the range for the random variable so that the desired output can be achieved.

→ Infinite population has a stationary probability distribution with a finite mean μ and finite variance σ^2.

→ Sample variable and time doesn't affect population distribution

→ Variables that meet all these conditions are called independently and identically distributed.

→ Central limit theorem must be invoked to rely upon normal distribution of infinite population.

→ Only then we can apply estimation method to that variable taken from infinite population

- A random variable is drawn from an infinite population that has a stationary probability distribution with finite mean (μ) & finite variance (σ^2).

- The theorem states that the sum of n i.i.d variables, drawn from a population that has a mean of μ and a variance of σ^2, is approximately distributed as a normal variable with a mean of $n\mu$ and a variance of $n\sigma^2$.

- Let x_i ($i = 1, 2 \ldots n$) be the n i.i.d random variables.

Then normal variate:

$$Z = \frac{\sum_{i=1}^{n} x_i - n\mu}{\sqrt{n}\,\sigma}$$

PAGE NO.:
DATE / /

Simulation Run statistics

- Consider a single server system in which the arrivals occur with a poisson distribution and the service time has an exponential distribution

- Suppose, the study objective is to measure the mean waiting time, defined as the time entities spend waiting to receive service and excluding the service time itself.

- This system is commonly denoted by M/M/1 which indicates; first, that the inter-arrival time is distributed exponentially; second that the service time is distributed exponentially and third that there is one server. The M stands for Markovian, which implies an exponential distribution.

- In ~~exponential~~ a simulation run, the simplest approach is to estimate mean waiting time by accumulating the waiting time of n successive entities and dividing by n.

- This measure, the sample mean, is denoted $x(n)$ to emphasize the fact that its value depends upon the number of observation taken.

- If x_i ($i = 1, 2, \dots n$) are the individual waiting times (including the value 0 for those entities that do not have to wait), then :

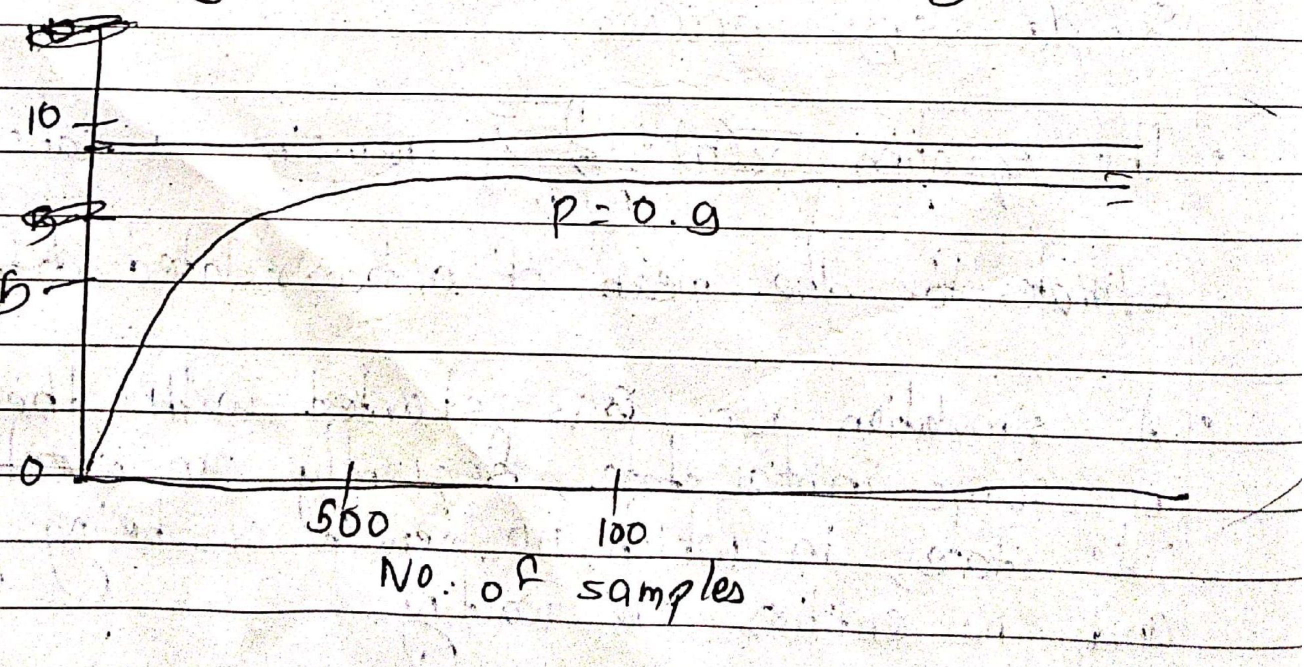
$$\bar{X}(n) = \frac{1}{n} \sum_{i=1}^{n} x_i$$

- Whenever a waiting line forms, the waiting time of each entity on the line clearly depends upon the waiting time of its predecessors.
- Any series of data that has this property of having one value affect other values is said to be autocorrelated.
- The sample mean of autocorrelated data can be shown to approximate a normal distubition as the Sample size increase.
- The equation $\bar{x}(n) = \frac{1}{n} \sum_{i=n}^{n} x_i$ remains a satisfactory estimate for the mean of autocorrelated data.
- A simulation run is started with the system in some initial state, frequently in ~~ideal state~~ idle state, in which no service is being given and no entities are waiting.

- The early arrivals then have a more than normal probability of obtaining service quickly, so a sample mean that includes the early arrivals are biased.

- For a given sample size starting from a given initial condition, the sample mean distribution is stationary; but, if the distributions could be compared for different sample size, the distribution would be slightly different.

- The following figure is based on theoretical results, which shows how the expected value of sample mean depends upon the sample length, for the M/M/1 system, starting from an initial empty state, with a server utilization of 0.9.



PAGE NO.:
DATE: / /

- ## Replications of Runs

→ The precision of results of dynamic stochastic can be increased by repeating the experiment with different random number strings.

- For each replication of a small sample size, the sample mean is determined.

- The sample means of the independent runs can be further used to estimate the variance of distribution. Let X_{ij} be the i^{th} observation in j^{th} run, then the sample mean and variance for the j^{th} run are:

$$\overline{X_j(n)} = \frac{1}{n} \sum_{i=1}^{n} X_{ij}$$

$$S_j^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left[X_{ij} - \overline{X_j(n)}\right]^2$$

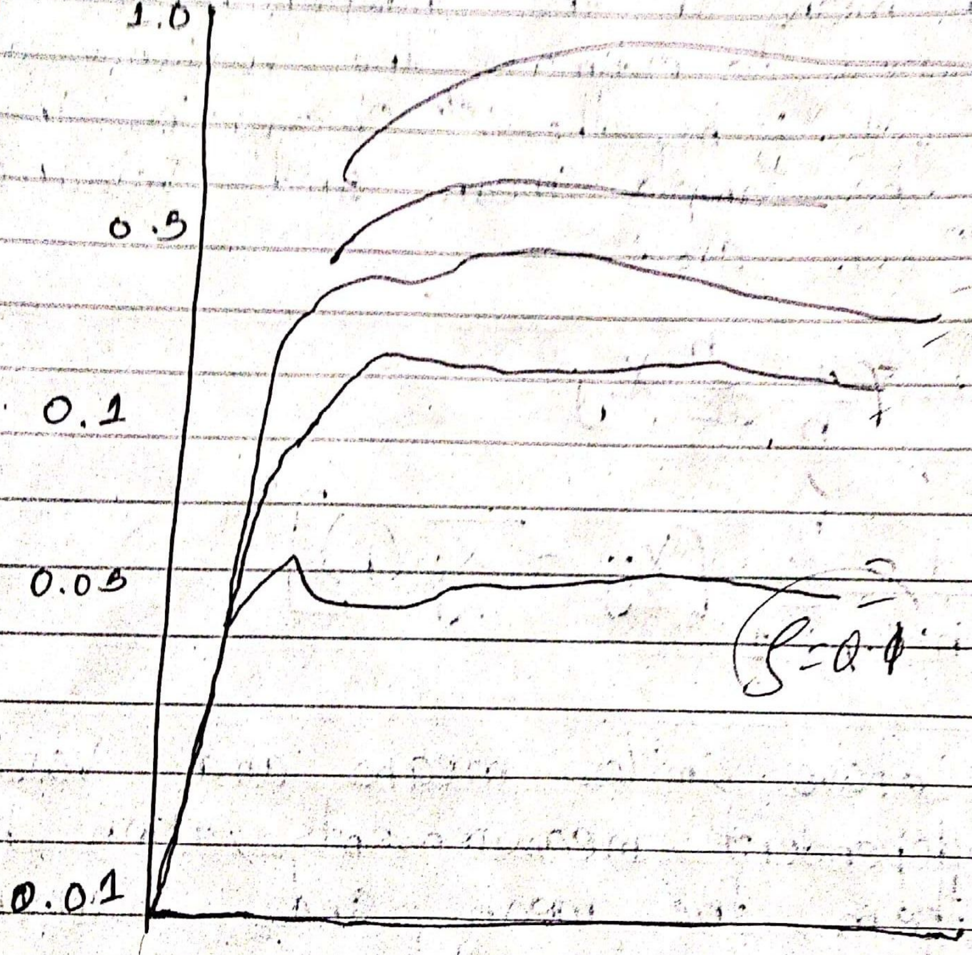
- When we have similar means and variances for m independent measurements, then by combining them, the mean and variance for the population can be obtained as:

PAGE NO:
DATE: / /

$$\bar{X} = \frac{1}{P}\sum_{j=1}^{P} \bar{x}_j(n)$$

$$S^2 = \frac{1}{P}\sum_{j=1}^{P} s_j^2(n)$$

- The following figure shows the result of applying the procedure to experimental results for the M/M/1 system.



Scanned with CamScanner

Elimination of Initial Bias

→ Two general approaches can be taken to remove the bias:

→ the system can be started in a more representative state that the empty state

→ or the first part of the simulation can be ignored.

→ The ideal situation is to know the steady state distribution for the system, and select the initial condition from that distribution.

→ In this study previously discussed, repeated the experiments on the M/M/1 system, supplying an initial waiting line for each run, selected at random from the known steady state distribution of waiting line:

→ The case of 40 repetitions of 320 samples, which prieviously resulted in a coverage of only 9% was improved to coverage of 88%.

PAGE NO.:
DATE: / /

→ The more common approach to removing the initial bias is to estimate an initial section of the run.

→ The run is started from an idle state & stopped after a certain period of time.

→ The run is then restarted with statistics being gathered from the restart.

→ It is usual to program the simulation so that the statistics are gathered from the beginning and simply wiped out the statistics gathered up to the poin of restart.

→ No simple rules can be given to decide how long an interval should be eliminated.

→ The disadvantage of eliminating the first part of a simulation run is that the estimate of the variance, needed to establish a confidence limit, must be based on less information.

→ The reduction in bias, ∴ is obtained at the price of increasing the confidence interval size.

PAGE NO.:
DATE: / /

chapter - 5

- **Random Number generation Techniques:-**

1) Linear Congruential Generator (LCG)

$$z_i = (a z_{i-1} + c) \% M$$

$\because$ a = Multiplicative factor
c = Incremental factor
z_{i-1} = Seed value
M = Modulus value

Multiplicative ma $C = 0$
Incremental ma $a = 1$

$$u_1 = \frac{z_i}{M}$$

PAGE NO.:
DATE: / /

- Test of Random Number

1) Frequency Test - Uniformity

i) k-S Test

→ The procedure used to perform k-S test are as follows:-

Step 1: Rank the random number from smaller to larger i.e

$$R_{(1)} \leq R_{(2)} \leq \dots R_{(n)}$$

Step2: Compute

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

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

Step 3: Compute

$$D = \max(D^+, D^-)$$

step4: for specified level of specificance α and given sample size N, take the table value D_α.

Step 5: Compare the table value D_α with D. If $D \leq D_\alpha$, the null hypothesis can't be registered i.e the numbers are uniformly distributed.

PAGE NO.:
DATE /

ii) Chi-Square

- Divide the sample into n intervals of equal size (0-0.1, 0.11-0.2, . . . 0.91-1.0)
- Count the observed frequency on the given range
- Calculate the estimated frequency using $E(i) = \frac{N}{n}$, ∵ N => total no. of observation, n => Number of intervals.
- Perform

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

- Generate critical value, $\chi^2_{\alpha, n-1}$ from the table for level of significance α
- If $\chi_0^2 \leq \chi^2_{\alpha, n-1}$, the null hypothesis that the numbers are uniformity distributed can't be rejected.

PAGE NO.:
DATE /

- Independence Test

1) Run Test
   - Run up & down
   - Run above & below mean
   - length of run

i) Run up and down Test

a = Total no. of + and - sequence line
(Total no. of runs in a truly random sequence)

$\mu_0 = \frac{2N-1}{3}$ (∵ N ⇒ total no. of observation)

$\sigma_a^2 = \frac{16N-29}{90}$

So,

$Z_0 = \frac{a - \mu_0}{\sigma_a}$ $\left[∵ \frac{a - \left[\frac{2N-1}{3}\right]}{\sqrt{\frac{16N-29}{90}}} \right]$

$Z_0 < Z_{\alpha/2}$

PAGE NO.:
DATE. / /

ii) Run above and below Mean

→ Here a run can be achieved by comparing the random numbers by the mean value (0.495)

→ Number larger than mean represented by '+'.
→ " less than " " by '−'

$n_1 \Rightarrow (+)$ b = total no. of sequence
$n_2 \Rightarrow (-)$

$$\mu_a = \frac{2 n_1 n_2}{N} + \frac{1}{2} \quad \text{— (1)}$$

$$\sigma_b^2 = \frac{2 n_1 n_2 (2 n_1 n_2 - N)}{N^2 (N-1)} \quad \text{— (11)}$$

So,

$$Z_0 = \frac{b - \mu_b}{\sigma_b}$$

iii) length of runs ~~above~~ up & down

$$E(Y_i) = \frac{2}{(i+3)!}\left[N(i^2+3i-1)-(i^3+3i^2-i-4)\right], \quad i \leq n-2$$

check note: $i \Rightarrow +, -$ ~~sequence count garne~~, check notes

$O_i \Rightarrow +, -$, sequence count garne

Same as k-S test

i	O_i	$E(Y_i)$	$[O_i - E(Y_i)]^2$	$(O_i - E(Y_i))^2 / E(Y_i)$

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

PAGE NO.:
DATE. / /

- length of run above and below mean

$$E(Y_i) = \frac{Nw_i}{E(I)}, N > 20$$

for w_i:

$$w_i = \left(\frac{n_1}{N}\right)^i \left(\frac{n_2}{N}\right) + \left(\frac{n_1}{N}\right)\left(\frac{n_2}{N}\right)^i$$

for $E(I)$

$$E(I) = \frac{n_1}{n_2} + \frac{n_2}{n_1}$$

Then

$$\chi_0^2 = \sum_{i=1}^{L} \frac{[O_i - E(Y_i)]^2}{E(Y_i)}$$

chapter - 2

- Monte Carlo Simulation

→ It is a technique of experimental sampling with random number or method of trial which can be used to solve many problems which are otherwise difficult or sometimes impossible.

→ The method can't be used to attempt high accuracy and hence it is suitable for those problems that doesn't require high degree of accuracy.

→ They are stocastic technique and make use of random number and probability to solve the problem.

→ It can be used solve both stocastic and deterministic problems. when we solve the deterministic problem using montecarlo we first convert the deterministic model to stocastic model. Eg- Calculating the value of pi (π).

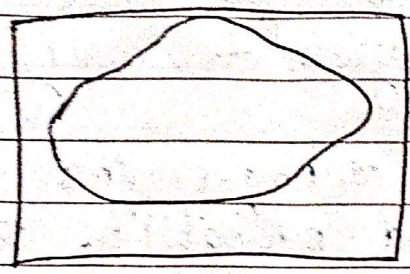
- Some eg:
  i) To find the area of irregular surfaces
  ii) Numerical integration.
  iii) Random walk problem

i) Finding the area of irregular surface.

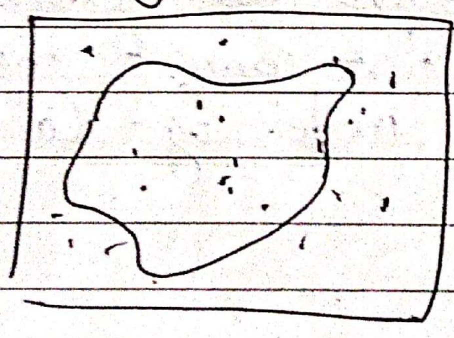
→ Let us consider an irregular shape as shown in

fig: Irregular shape

Enclose the shape with rectangle as below



Inside rectangle mark random 'N' number of dots as shown in figure.



Let there be 'M' dots inside the irregular surface. and 'N' be the total no. of dots inside rectangle.

PAGE NO.:
DATE /

Count the number of dots inside irregular figure. ~~let 'M~~

If 'F' be the area of irregular fig then, we can write

$$\frac{F}{A} = \frac{M}{N}$$

where, F = Area of irregular fig
A = Area of rect
M = No. of dots inside irregular
N = " " " " rect.

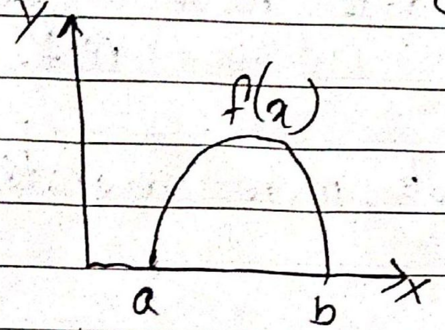
→ large the numbers of dots, more accurate will be the result.

Scanned with CamScanner

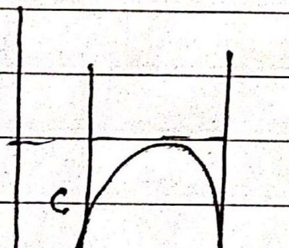
• Numerical integration

→ The integration of single variable over a given range gives the area under graph representing the function procedure.

→ Let us consider a function $f(x)$ having lower & upper bound a & b respectively.



Now, draw a rectangle. The side of a rect are (b-c) and c



Area of rect $= (b-a) \times c$

Now, mark random number of dots inside the rectangle and count dots let 'M' be the dots inside ~~cure~~ curve & 'N' be total dots

Scanned with CamScanner

PAGE NO.:
DATE

Then,

$$\frac{\text{Area of irregular surface}}{\text{Area of rect}} = \frac{M}{N}$$

i.e

$$\frac{\text{Area of curve}(f)}{\text{Area of rect}} = \frac{M}{N}$$

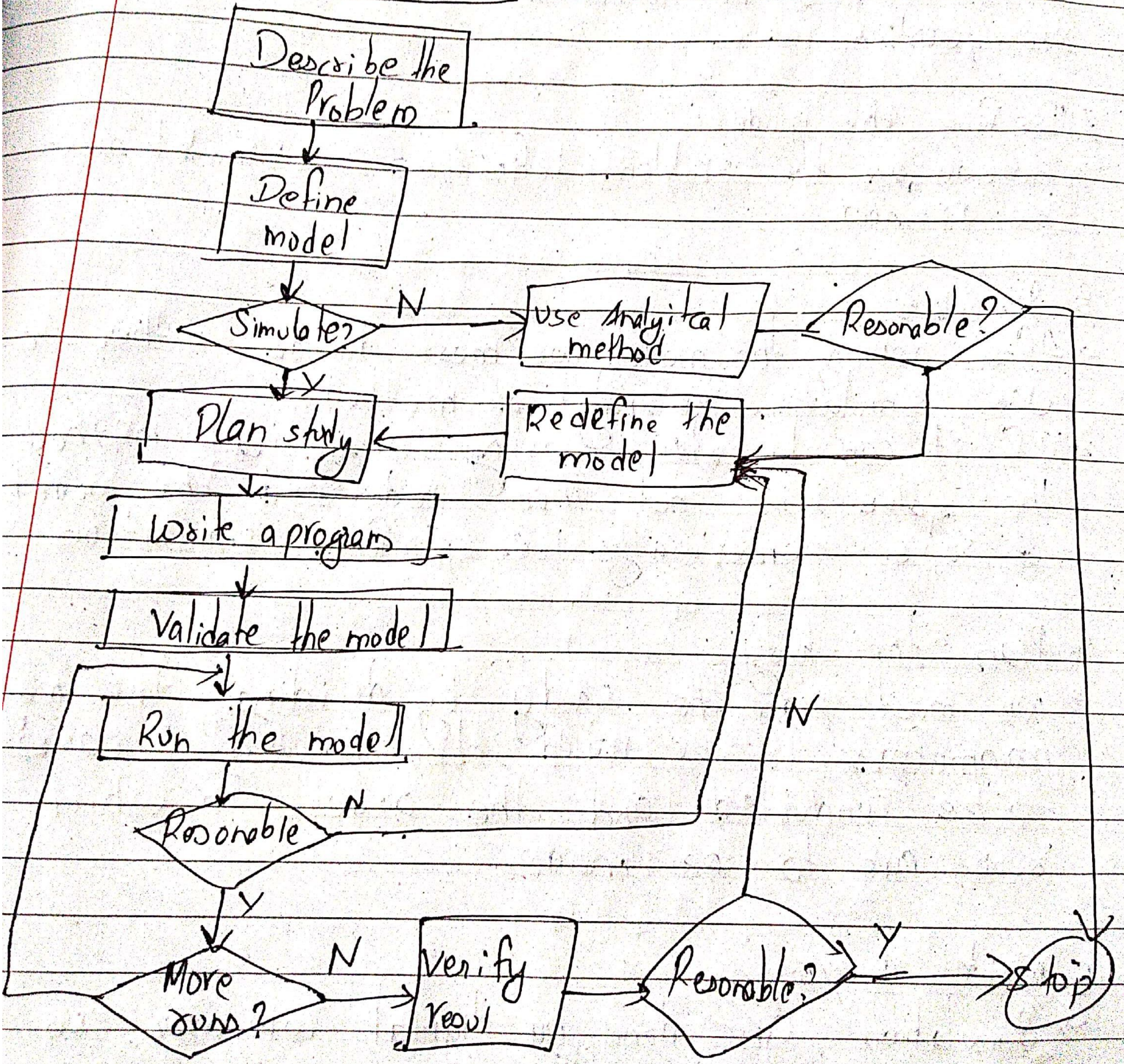
$$\boxed{\text{i.e} \quad \int_a^b f(x)\,dx = \frac{M}{N} \times (b-a) \times c}$$

→ The accuracy improves as number a 'N' increase

- Experimental nature of simulation

check notes

- steps of simulation



Step 1 : Define the problem

→ An initial setup is to describe problem to be solved in concized manner. The description must be enough to answer questions asked & what need to be taken in order to answer the question.

2. Define the model
→ Based on the problem definition, a model must be defined

3. Simulate
→ After defining the model, we must decide either to use simulation or analytical method.
→ If it can be solved using analytical method the solve, if not then redefined the model & solve numerically. i.e perform simulation

4. Plan the study
→ we must plan the study by deciding the major parameters to be varied, the no. of cases to be conducted and the order in which the runs are to be made.

5. Write a program
→ Simulation are done by digital computer, so program must be written.

6. Validate the model
→ The model must be valid before the begining of major set of runs.

7. Run the ~~pro~~ model
→ The stage will then move into the start of executing a series of runs according to study plan. It is essential to repeat the run with different set of random number so that more than one sample result is available

8 Verify result.
→ After making multiple runs, we get various samples & those samples must be verified. If the approximation is found, stop the process otherwise redifine the model & repeat

• Types of System Simulation

1) Numerical computation technique of discrete System
2) " " " of Continuous "

• Distributed Lag Model

→ Models that have the properties of changing only at fix interval of time and basing current values of the variables on the basis of current values & the values that occur in previous interval are called distributed lag models.

→ These models are used for economic studies.

→ They represent continous system but the data is available at fix interval of time.

let C = Consumption
I = investment
T = Tax
G = Government expenditure
Y = National income.

$$C = 20 + 0.7(Y_{-1} - T_{-1})$$
$$I = 2 + 0.1Y_{-1}$$
$$T = 0.2Y_{-1}$$
$$Y = C_{-1} + I_{-1} + G_{-1}$$

PAGE NO.:
DATE

To solve;

$I = 2 + 0.2Y_{-1}$

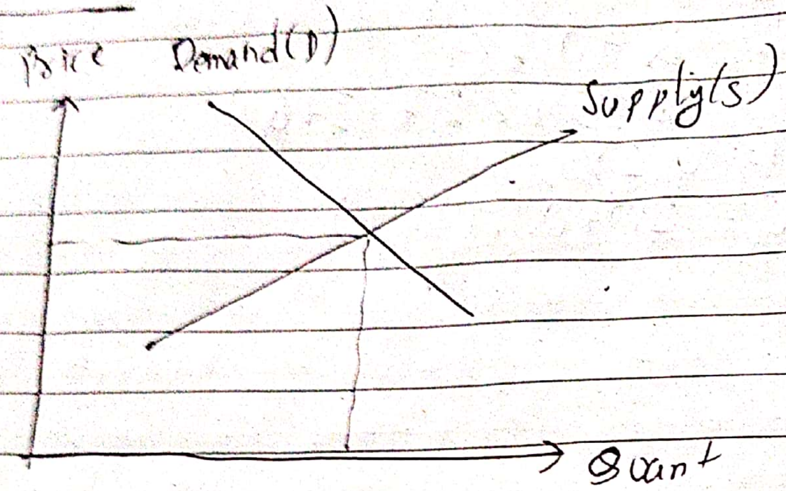
$Y = 45.45 + 2.27(I+G)$

$T = 0.2Y$

$C = 20 + 0.7(Y-T)$

PAGE NO.:
DATE. / /

• Cobweb Model



⇒ Consider the static mathematical model for market, demand and supply model, we have

$$D = a - bp$$
$$S = c + dp$$
$$D = S$$

⇒ Generally, supply should be dependent on demand from previous marketing period.

• But the demand however will respond to current price. So, writing the market model in distributed lag model we get.

$$D = a - bp$$
$$S = c + dp_{-1}$$
$$D = S$$

• Time advance mechanism / Check notes.

→ Simulation clock is a variable in a simulation model that gives the current value of simulated time. There is no particular unit for simulation clock time.

PAGE NO.:
DATE. /

Chapter-4

Discrete System Simulation

• Components and Organization of discrete System

1. System state

→ The collection of state variables necessary to describe the system at a particular time.

2) Simulation clock
→ A variable giving the current value of simulated time

3) Event List
→ A list containing the next time when each type of event will occur.

4) Statistical Counters
→ Variables used for storing statistical information about System performance.

5) Initialization routine
→ A sub ~~routine~~ program that initializes the simulation model at time 0.

PAGE NO.:
DATE / /

6) Timing Routine

→ A sub-routine that determines the next event from the event list and then advances the simulation clock to the the time when the event is to occur.

7) Event Routine

→ A sub-program that updates the system state when a particular type of event occur.

8) Library Routines

→ A set of sub-programs used to generate random observation that were determined as part of simulation model.

9) Report Generator

→ A sub-program that computer estimates of the desired measures of performance and ~~procedures~~ produces a report when the simulation ends.

10. Main Program

→ A sub program that invokes the timing routine to determine the next event and then transfer control to corresponding event routine to update the system state appropriately.

PAGE NO.:
DATE. / /

• ~~Simulation Time / Time Advance Mechanism~~

→ ~~Discrete event simulation concerns~~

• Generation of Arrival Patterns.

→ Arrival patterns for particular system is specified for simulation. The exogeneous arrivals can be designed for simulation. Two basic arrival patterns are:-

i) Trace Driven simulation

→ It refers to the process of gathering a sequence of inputs based on the observation of a real system.

→ It is an important tool in many simulation applications in which the model's input are derived from a sequence of observations made on a real system.

→ When there is no interaction between exogenous arrivals and the endogeneous events of a system it is permissible to create a ~~simulation~~ sequence of arrival in preparation for the simulation.

→ Usually the simulation proceeds by creating new arrivals as they needed.

PAGE:
DATE:

→ Here, program monitors can be attached to the running system to extract data with no or little disturbance to running system.

ii) Bootstrapping method:

→ It usually refers to a self-starting process that is supposed to proceed without external input.

→ It is the process of making one entity creates its successor.

→ Here, the arrival time of next entity is recorded, and when the clock reaches this event time, the event of entering the entity into the system is executed, and the arrival of the following entity is immediately calculated from the inter-arrival time distribution.