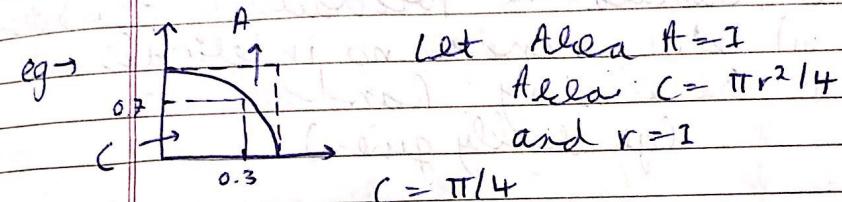


MODELLING AND SIMULATION

PAGE NO. _____
DATE : _____

MONTE CARLO SIMULATION
A method of estimating the value of an unknown quantity using the principles of INFERENTIAL STATISTICS

→ Variance should be minimum in parent sample for this to be applicable



A = parent set
assume that $C = 75\%$ of A

$$\frac{\pi}{4} = \frac{75}{100} \quad \pi = 3.0$$

* total no. of points in A =

If $n_A = 10,000$, and $P = \text{no. of points in area C}$

$$C = \frac{P}{10000} \times A = \frac{P}{10000} =$$

$$\Rightarrow \frac{\pi}{4} = \frac{P}{10000}$$

$$\pi = \frac{4P}{10000}$$

PAGE NO. _____
DATE : _____

PAGE NO. _____
DATE : _____

Scheduled arrival time, starting at 8 am
Use following random no's for handling
above problem.

40, 82, 11, 34, 25, 66, 17, 79.

CATEGORY	TIME REQ.	NO. OF PATIENTS
Filling	45 min	40
Cream	60	15
Cleaning	15	15
Extracting	45	10
Checkup	15	20
		100

ANS.

1. Category	Time Req.	Prob.	CP	R. No Interval
Filling	45	0.40	0.40	0-39
Cream	60	0.15	0.55	40-54
Cleaning	15	0.15	0.70	55-69
Extracting	45	0.10	0.80	70-79
Checkup	15	0.20	1.00	80-99

2. R. No.	Patient	Category	Time Req.
40	1	Cream	60
82	2	Checkup	15
11	3	Filling	45
34	4	Filling	45
25	5	Filling	45
66	6	Cleaning	15
17	7	Filling	45
79	8	Extracting	45

STEPS IN MONTE CARLO

- (i) Establishing probability distribution (given)
- (ii) Cumulative probability distribution
- (iii) Setting random no. intervals
- (iv) Generating Random no. (generally given)

(1) Strong is a dentist who schedules all of his patients for 30 min appointment. Some of the patients take less or more 30 min depending on the type of dental work to be done.

The following summary shows various categories of work, time needed to complete the work, and their probabilities.

Simulate the dentist clinic for 4 hrs & find out avg waiting time for patient as well as idleness for the doctor.

Assume that all patients show up at the clinic at exactly on their

PAGE NO. _____
DATE : _____

PAGE NO. _____
DATE : _____

	Arrival	Service Start	Service End	Waiting (min)	Idle
1	8:00	8:00	9:00	0	0
2	8:30	9:00	9:15	30	0
3	9:00	9:15	10:00	15	0
4	9:30	10:00	10:45	30	0
5	10:00	10:45	11:30	45	0
6	10:30	11:30	11:45	60	0
7	11:00	11:45	12:30	45	0
8	11:30	12:30	1:15	60	0
				285	0

$$\text{Avg waiting} = \frac{285}{8} = 35.62 \text{ min}$$

idle time = 0 min

(1)	DAILY DEMAND	PROB
	0	0.01
	15	0.15
	25	0.20
	35	0.50
	45	0.12
	50	0.02

Check Simulation for next 10

days
cakes mode = 35/day

Avg daily Demand = ?

R N - 40, 52, 11, 34, 66, 17, 79 Cumulative Prob

ANS.	DAILY DEMAND	PROB	C P	REGULAR INTERVAL
	0	0.01	0.01	00
GIVEN	15	0.15	0.16	1-15
	25	0.20	0.36	16-35
	35	0.50	0.86	36-85
	45	0.12	0.98	86-97
	50	0.02	1.00	98-99

Day	R NO	Demand	35/day	Stock
1	48	35	35	0
2	78	35	35	0
3	09	15	35	20
4	51	35	35	0
5	56	35	35	0
6	77	35	35	0
7	15	15	35	20
8	14	15	35	20
9	68	35	35	0

R No belongs to which regular no. interval based on daily demand of that interval

absolute diff
of

PAGE NO. _____	DATE : _____
----------------	--------------

$$\begin{array}{|c|c|c|c|c|} \hline 10 & 9 & 15 & 35 & 20 \\ \hline \end{array}$$

$$\text{Avg demand} = \frac{\sum \text{Demand}}{10} = \frac{270}{10} = 27$$

PAGE NO. _____	DATE : _____
----------------	--------------

$$\begin{array}{|c|c|} \hline 153 & 0.12 \\ \hline 154 & 0.08 \\ \hline \end{array}$$

(2) The automobile Company manufactures around 150 scooters daily. Production varies from 146 - 150. The finished scooter are transported in containers containing 150 scooters. Use following random no. to find answer.
 RN: 80, 81, 76, 75, 64, 43, 18, 26, 10, 12, 65, 58, 69, 61, 57

Simulating no. of scooters waiting in a factory.

(i) Avg no. of empty space in containers.

ANS.	Production day	Prob	CP	Regular Interval
	146	0.04	0.04	00 - 03
	147	0.09	0.13	04 - 07
	148	0.12	0.25	13 - 24
	149	0.14	0.39	25 - 38
	150	0.11	0.50	39 - 49
	151	0.10	0.60	50 - 59
	152	0.20	0.80	60 - 79
	153	0.12	0.92	80 - 91
	154	0.08	1.00	92 - 99

Day	R.N.	Demand	150/day	Waiting Count
1	80	153	150	3
2	81	153	150	3
3	76	152	150	2
4	75	152	150	2
5	64	152	150	2
6	43	150	150	0
7	18	148	150	-2
8	26	149	150	-1
9	10	147	150	-3
10	12	147	150	-3
11	65	152	150	2
12	58	151	150	1

Production day	Prob
146	0.04
147	0.09
148	0.12
149	0.14
150	0.11
151	0.10
152	0.20

13	69	152	150	2
14	61	152	150	2
15	57	151	150	1

avg waiting scooties = sum of positive no's = $\frac{3+3+2+2+2+2+1+2+2+1}{15} = \frac{20}{15} = \frac{4}{3}$

avg empty space = sum of negative no's = $\frac{2+1+3+3}{15} = \frac{9}{15} = \frac{3}{5}$

→ EXPECTED OR MEAN VALUE

$$E(x) = \sum_{\text{all } i} x_i p(x_i) \text{ for discrete } x_i$$

$$E(x) = \int_{-\infty}^{\infty} x f(x) dx \text{ for continuous}$$

↳ PDF

- * Variance → always gives +ve value
- a Standard deviation → gives a range i.e. -ve to +ve.
- * Standard deviation which we find is always corresponding to a particular point in that range that's why +ve

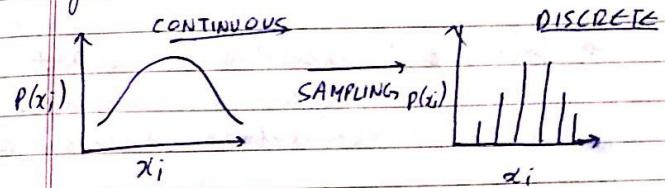
→ RANDOM VARIABLE

↓
DISCRETE
CONTINUOUS

→ R_x (Range of x) of a RANDOM VARIABLE is finite numbers or countably infinite for discrete random variable

PROPERTIES:

- (i) $p(x_i) \rightarrow$ probability that x takes x_i
- (ii) $p(x_i) \geq 0 ; \sum_{\text{all } i} p(x_i) = 1$
- (iii) $(x_i, p(x_i))$ is called prob distribution func.
- (iv) $p(x_i)$ is called probability mass func.



CONTINUOUS FUNCTION

e.g. → Dice tossing experiment
 $x_i = \text{no. of spots on up face}$
 $R_x = \{1, 2, 3, 4, 5, 6, 7\}$

Assume prob of a given face getting landed is proportional to no. of spots shown.

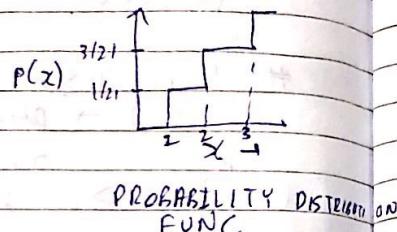
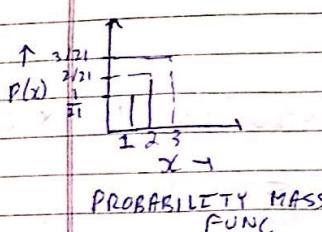
PAGE NO. _____
DATE : _____

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

If prediction to be made after life 2-3 years

PAGE NO. _____
DATE : _____

ANS.	x_i	1	2	3	4	5	6
	$P(x_i)$	$\frac{1}{21}$	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{6}{21}$



CONTINUOUS RANDOM VARIABLE

If R_x of a random variable is an interval or collection of intervals, x is continuous Random variable

→ Collection: collection of height
collection of space
collection of face cards.

→ Prob that x lie in interval (a, b) is

$$P(a \leq x \leq b) = \int_a^b f(x) dx.$$

RANDOM VARIABLE

(i) $f(x) \geq 0 \quad \forall x \in R_x$

(ii) $\int_{R_x} f(x) dx = 1$

(iii) $f(x) = 0 \rightarrow$ when x is not in R_x

$$\text{ANS. } \int_0^{3/2} \frac{1}{2} e^{-x/2} dx = \frac{1}{2} \left[e^{-x/2} \right]_0^{3/2} = e^{-3/2} - e^{-2/2} = 0.145$$

CUMMULATIVE DISTRIBUTION FUNC

(i) If x is discrete:
 $CDF = \sum_{x_i \in X} P(x_i)$

(ii) If x is continuous:

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

* If F is a non-decay function
then $\lim_{x \rightarrow \infty} F(x) = 1$

* If F is a decay function
then $\lim_{x \rightarrow \infty} F(x) = 0$

VARIANCE

$$\text{V} = E [x - E[x]]^2$$

$$\text{V} = E[x^2] - (E(x))^2$$

BERNOULLI DISTRIBUTION

- * useful for independent cases
- * For experiments consisting of n trials
- * each can be a success or a failure
- $x_j = 1$ If j th experiment is success and
 $x_j = 0$ If it is a failure.
- If n Bernoulli trials are independent
It is called Bernoulli Process distribution

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

$$\text{Mean} = E[x_j] = 0 \times q + 1 \times p = p$$

$$\text{Variance} = [0^2 q + 1^2 p] - p^2 = pq$$

$$= p(1-p) = pq$$

RANDOM NO.:

- * Eg OTP, captha
- * Every random no. generator is stochastic process

- * Disadvantages of R. No generator like coin-toss
 - (i) limited outcome
 - (ii) limited no's
 - (iii) storage required

FREQUENCY TEST →
INDEPENDENT TEST →
occurrence of random no. should be independent of other random no.

- * Every stochastic process contains random no.
eg, process = dice through thrown random no. = result (1, 2, 6)

- * CONSIDERATION:
 - Generation should be fast
 - It should be portable
 - Should have a sufficiently long cycle: it should generate random no. for many numbers
 - Should be replicable
we may generate same no again
 - Independent
 - Uniform

- * PSEUDO-RANDOM NO:
 - It should try to achieve properties of random no.

PAGE NO. _____
DATE: _____

CHALLENGES IN GENERATING PSEUDO R. NO

- Generation should be uniformly distributed
- mean of generated no. Should not lie too high or low
- variance of generated number should not be too high or too low
- Autocorrelation b/w no's (Should not follow any pattern)
- MED-SQUARE METHOD is an arithmetic method, i.e. which follows a basic calculation but not some formula

TECHNIQUES TO GENERATE PSEUDO RANDOM NO.

(i) Congruential generator (CG)

- Linear CG
- Mixed CG
- Multiplicative CG
- Combined linear CG

(ii) Linear CG

$$Z_i = (Z_{i-1} + c) \bmod M$$

Random no. $\rightarrow Z_0$ (From 0-1)
 $\frac{m}{m}$

Z_0 = seed value

1. Parameters $M=63, a=22, c=4, Z_0=19$

$$Z_i = [22Z_{i-1} + 4] \bmod 63$$

ANS.	i	$22Z_{i-1} + 4$	Z_i	R. NO
0			19	-
1		$22(19) + 4 = 422$	44	0.6984
:				

63	334	19	0.3016
64	422	44	0.2354

↳ repetition after M iterations

→ CHI-SQUARE TEST → for large data

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad O_i = \text{observed no.} \quad E_i = \text{Expected no.}$$

From uniform distribution
no in each class $\left[\frac{N}{n} = E_i \right]$
↳ SUMMATION

* $\chi^2 > \chi^2_{\alpha} \rightarrow \text{UNIFORM}$
↳ GIVEN

* $\chi^2 < \chi^2_{\alpha} \rightarrow \text{NON-UNIFORM}$

→ CALCULATE EXPECTED FREQ 2x2 TABLE

A/B	B ₁	B ₂	Total
A ₁	F ₁₁	F ₁₂	N ₁
A ₂	F ₂₁	F ₂₂	N ₂
Total	N ₃	N ₄	N

m × n table

$$v = \text{degree of freedom} \\ = (m-1)(n-1)$$

$$f_{e11} = \frac{N_1 \times N_3}{N} \quad f_{e12} = \frac{N_1 \times N_4}{N}$$

$$f_{e21} = \frac{N_2 \times N_3}{N} \quad f_{e22} = \frac{N_2 \times N_4}{N}$$

- (1) In a sample survey of public opinion answer to the question
 (i) Do you drink
 (2) Are you in favour of local opinion sale liquor are tabulated below

Ques-1

	Yes	No	Total
ques 2	56	31	87
No	18	6	24
Total	74	37	111

can you infer or not the local

opinion on sale of liquor is dependent on individual drink.
 Given that $\chi^2 = 3.841$ corresponding to $\beta_{0.1} = \alpha$

$$f_{011} = \frac{87 \times 24}{111} \quad f_{012} =$$

ANS. H_0 = hypothesis that sale of liquor is dependent on individual drink.

$$f_{011} = \frac{87 \times 74}{111} = 58$$

$$f_{011} = 56$$

$$f_{012} = \frac{87 \times 37}{111} = 29$$

$$f_{012} = 31$$

$$f_{021} = \frac{24 \times 74}{111} = 16$$

$$f_{021} = 18$$

$$f_{022} = \frac{24 \times 37}{111} = 8$$

$$f_{022} = 6$$

$$\chi^2 = \frac{58 - 56}{58 + 56} + \frac{31 - 29}{31 + 29} + \frac{16 - 18}{16 + 18} + \frac{8 - 6}{8 + 6}$$

$$\chi^2 = \frac{4}{58} + \frac{4}{29} + \frac{4}{16} + \frac{4}{8}$$

$$= 0.068 <$$

$$0.9568 < 3.841$$

\therefore IT IS UNIFORM

AUTO-CORRELATION TEST

- (i) concerned with dependency b/w no's in sequence
- (ii) Computation of auto-correlation b/w every m numbers (m is known as "lag") starting with i th no.

(iii) δ_{im} b/w numbers

↳ lag $R_i, R_{i+m}, R_{i+2m}, \dots, R_{i+(M+1)m}$

↳ starting point

↓ highest integer

is to be found

$M = \text{largest integer s.t.}$

$$i + (M+1)m \leq N$$

- (iv) A non-zero auto-correlation implies lack of dependence, i.e.

PAGE NO. _____
DATE : _____

PAGE NO. _____
DATE : _____

following detail test is appropriate

$H_0: \delta_{im} = 0 \rightarrow$ Always dependent
 $H_1: \delta_{im} \neq 0 \rightarrow$ may or may not be independent

$\rightarrow \delta_{im} \neq 0$

$$Z_0 = \frac{\delta_{im}}{\sigma_{\delta_{im}}} \quad Z_{x/2} = \text{given in graph}$$

given in graph $-Z_{x/2} \leq Z_0 \leq Z_{x/2}$

$$S = \frac{1}{M+1} \left[\sum_{k=0}^M R_{i+k} R_{i+(k+1)m} \right] - 0.25$$

$$\sigma_{\delta_{im}} = \sqrt{\frac{13M+7}{12(M+1)}}$$

- Q. Starting from second, check auto correlation for 7th and 12th given 40 Random no.'s

ANS. $N = 40 \quad m = 12 - 7 = 5 \quad i = 2$

$$i + (M+1)m \leq N$$

$$2 + (M+1)5 \leq 40$$

$$2 + 5M + 5 \leq 40$$

$$5M \leq 33$$

$$M \leq 6 \quad M = 6$$

$$S = \frac{1}{7} [R_{21} R_7 + R_7 R_{12} + R_{12} R_{17} + \\ R_{17} R_{22} + R_{22} R_{27} + R_{27} R_{32} + R_{32} R_{37}] - 0.25$$

$$\sigma = \sqrt{\frac{13 \times 6 + 7}{12(7)}}$$

Now compare $\text{Z value and } \frac{S}{\sigma}$

- If $S > 0 \rightarrow$ +ve auto-correlation
- If $S < 0 \rightarrow$ -ve auto-correlation

- * +ve auto-correlation means high random no's are followed by high and low by low
- * -ve auto-correlation means low random no's are followed by high and high by low

- (1) Test 3rd, 8th, 13th and so on in sequence are auto-correlated using $d = 0.05$

0.12, 0.01, 0.23, 0.28, 0.89, 0.31,
 0.64, 0.28, 0.83, 0.93, 0.99, 0.15
 0.33, 0.35, 0.91, 0.41, 0.60, 0.77
 0.75, 0.88, 0.68, 0.49, 0.05, 0.43
 0.95, 0.58, 0.19, 0.36, 0.69, 0.87

