UNIT-I

Probability & Random variable

Short Answer questions

(1) Define probability, Conditional probability & Bayes Theorem

Defination of probability:

In a Random Experiment, Let there be an 'n' mutually exclusive and equally likely elementary events. Let E be an event of Experiment. The probability of E

$$P(E) = \frac{m}{n} = \frac{\text{Number of elementary events in } E}{\text{Total number of elementary events in the Random experiment}}$$

Probability - Axiomatic approach:-

Def: Let 'S' be a finite Sample Space. A real valued function p from the power set of 'S' into R is called a probability function on 'S' if the following three axioms are satisfied

i) Axiom of positivity: $P(E) \geq 0$ for every subset E of S

(ii) Axiom of Certainity:- $P(S) = 1$

iii) Axiom of Union:- If E_1 and E_2 are disjoint subset of 'S' Then

$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

Conditional probability:-

Def:- If E_1 and E_2 are two events in a sample space 'S' and $p(E_1) \neq 0$, Then the probability of E_2, after the event E_1 has occurred, is called the conditional probability of the event of E_2 given E_1 and is denoted by $p(E_2/E_1)$

$$p(E_2/E_1) = \frac{p(E_1 \cap E_2)}{p(E_1)}$$

Similarly we define $p(E_1/E_2) = \frac{p(E_1 \cap E_2)}{p(E_2)}$

Bayes Theorem

Statment:-

$E_1, E_2 \ldots E_n$ are n mutually exclusive and exhaustive events such that $p(E_i) > 0$ $(i = 1, 2 \ldots n)$ in a sample space 'S' and A is any other event in S intersecting with every E_i (i.e A can only occur in combination with any one of the events $E_1, E_2 \ldots E_n$) such that $p(A) > 0$

if E_i is any of the events of $E_1, E_2 \ldots E_n$ where $p(E_1)\, p(E_2) \ldots p(E_n)$ and $p(A/E_1)\, p(A/E_2) \ldots p(A/E_n)$ are known then

$$p(E_k/A) = \frac{p(E_k) \cdot p(A/E_k)}{p(E_1)p(A/E_1) + p(E_2)p(A/E_2) + \cdots + p(E_n)p(A/E_n)}$$

(2) Define Random variable & Types of Random variables

Random variable :-

Def :- A real variable X whose value is determined by the outcome of a random experiment is called a random variable

Types of Random variable

i) Discrete Random variable

A random variable X which can take only a finite number of discrete values in an interval of domain is called a discrete Random variable

(ii) Continuous Random variable:

A Random variable 'X' which can take values continuously i.e which takes all possible values in a given Interval is called continuous random variable

(3) Define probability distribution function

Let 'x' be a random variable. Then the probability function associated with 'x' is defined as the probability that the outcome of an experiment will be one of the outcomes for which $X(s) \leq x, x \in R$

(4) A continous random variable has the p.d.f

$$f(x) = \begin{cases} k + \frac{x}{6} & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$ find 'k'

Sol:-

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\int_{0}^{3} (k + x/6)\,dx = 1$$

$$\left[kx + \frac{x^2}{12}\right]_0^3 = 1$$

$$3k + \frac{9}{12} = 1$$

$$3k = 1 - \frac{9}{12}$$

$$\cancel{3}k = \frac{\cancel{3}}{12} \Rightarrow k = \frac{1}{12}$$

(5) Throwing two dice, find the probability of getting sum is even

Throwing two die $6^2 = 36$ out comes $= n$

favorable out comes $m = 18$

$$\begin{bmatrix} (1,1) & (1,3) & (1,5) \\ (2,2) & (2,4) & (2,6) \\ (3,1) & (3,3) & (3,5) \\ (4,2) & (4,4) & (4,6) \\ (5,1) & (5,3) & (5,5) \\ (6,2) & (6,4) & (6,6) \end{bmatrix}$$

$$P(E) = \frac{18}{36} = \frac{1}{2}$$

(6) In a pack of Cards, find the probability of getting Ace (or) Spade

Sol:-

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(\text{Ace (or) Spade}) = P(\text{Ace}) + P(\text{Spade}) - P(\text{Ace} \cap \text{Spade})$$

$$= \frac{4C_1}{52C_1} + \frac{13C_1}{52C_1} - \frac{1C_1}{52C_1}$$

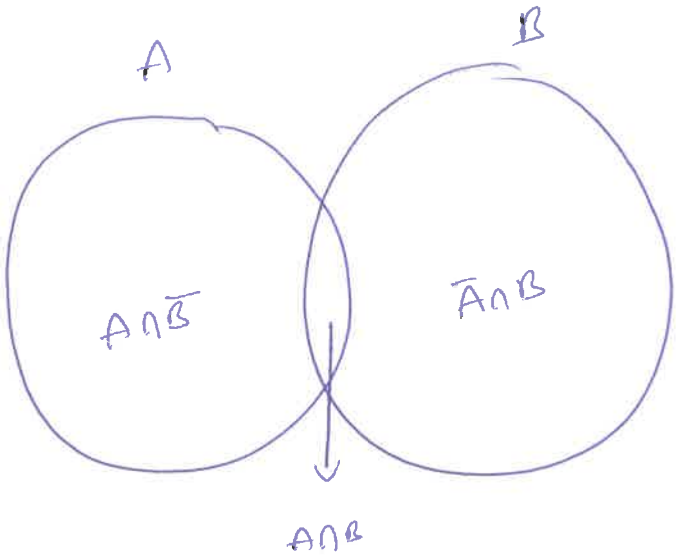
$$= \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52}$$

long Answer questions

(1) State and prove addition theorem on probability

Statment:- if S is a Sample space and A and B are any Two events

$$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



Using above diagram

$A \cup B = (A \cap \bar{B}) \cup (A \cap B) \cup (\bar{A} \cap B)$

A and B disjoint sets by using Axiomatic def

$P(A \cup B) = P((A \cap \bar{B}) \cup (A \cap B) \cup (\bar{A} \cap B))$

$P(A \cup B) = P(A \cap \bar{B}) + P(A \cap B) + P(\bar{A} \cap B) \rightarrow (1)$

$A = (A \cap \bar{B}) \cup (A \cap B)$

$P(A) = P(A \cap \bar{B}) + P(A \cap B)$

$P(A \cap \bar{B}) = P(A) - P(A \cap B) \rightarrow (2)$

$B = (\bar{A} \cap B) \cup (A \cap B)$

$P(B) = P(\bar{A} \cap B) + P(A \cap B)$

$P(\bar{A} \cap B) = P(B) - P(A \cap B) \rightarrow (3)$

(2) & (3) sub in (1)

$P(A \cup B) = P(A) - P(A \cap B) + \cancel{P(A \cap B)} + P(B) - P(A \cap B)$

so

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(2) Box 'A' Contain '5' red and '3' white marbles and box B contains '2' red and '6' white marbles. if marble is drawn from each box, what is the probability that they are both of same colour.

Sol:-

Suppose E_1 = the event that the marble is drawn from box 'A' and is red

$$P(E_1) = \frac{1}{2} \cdot \frac{5}{8} = \frac{5}{16}$$

and E_2 = The event that the marble from box 'B' and red

$$P(E_2) = \frac{1}{2} \cdot \frac{2}{8} = \frac{1}{8}$$

The probability that both the marbles are red is

$$P(E_1 \cap E_2) = P(E_1) \cdot P(E_2) = \frac{5}{16} \cdot \frac{1}{8} = \frac{5}{128}$$

Let E_3 = The event that the marble from box A and is white

$$P(E_3) = \frac{1}{2} \cdot \frac{3}{8} = \frac{3}{16}$$

Let E_4 = The event of marble from box B is white

$$P(E_4) = \frac{1}{2} \cdot \frac{6}{8} = \frac{3}{8}$$

and $P(E_3 \cap E_4) = \frac{3}{16} \cdot \frac{3}{8} = \frac{9}{128}$

The probability that the marbles are of same colour

$$= P(E_1 \cap E_2) + P(E_3 \cap E_4)$$

$$= \frac{5}{128} + \frac{9}{128} = \frac{14}{128} = \frac{7}{64} = 0.109$$

(3) Two factories produce identical clocks. The production of the first factory consists of 10,000 clocks of which 100 are defective. The second factory produces 20,000 clocks of which '300' are defective. What is the probability that a particular defective clock was produced in the first factory?

Sol: Two factories is denoted by A and B

Output produced by first factory A = 10,000

Output produced by second factory B = 20,000

probability that items produced by A are defective

$$P(D/A) = \frac{100}{10000} = 0.01$$

Similarly $P(D/B) = \frac{300}{20,000} = 0.015$

$$P(A) = \frac{10000}{30000} = \frac{1}{3} = 0.33$$

$$P(B) = \frac{20000}{30000} = \frac{2}{3} = 0.66$$

find

$P(A/D)$ = prob of defective clock was produced by first factory

$$P(A/D) = \frac{P(A) \times P(D/A)}{P(A) \times P(D/A) + P(B) P(D/B)}$$

$$= \frac{(0.33) \times (0.01)}{(0.33)(0.01) + (0.66) 0.015} =$$

4) $P(A) = \frac{2}{3}$ $\quad P(B) = \frac{1}{5}$

Then p.T $\frac{2}{15} \leq P(A \cap B) \leq \frac{1}{5}$

proof:

Using Conditional probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$\Rightarrow P(A) = \frac{P(A \cap B)}{P(B)}$$

$$P(A \cap B) = P(A)P(B)$$

$$P(A \cap B) = \frac{2}{3} \times \frac{1}{5}$$

$$P(A \cap B) = \frac{2}{15}$$

Since $\frac{2}{15} \leq \frac{1}{5}$ and $P(A \cap B) = \frac{2}{15}$

$$\Rightarrow \frac{2}{15} = P(A \cap B) \leq \frac{1}{5}$$

$$\frac{2}{15} \leq P(A \cap B) \leq \frac{1}{5}$$

5) From a lot of '10' items containing '3' defective. a sample of 4 items is drawn at Random. Let the random variable X denote the number of defective items in the sample. Find the probability distribution of 'X' when the sample without Replacement

Sol:- obviously X takes the Values 0, 1, 2 or 3

Given total No of items = 10

No of good items = 7

No of defective items = 3

$$P(X=0) = P(\text{no defective}) = \frac{^7C_4}{^{10}C_4} = \frac{7!}{4!3!} \times \frac{4!6!}{10!} = \frac{1}{6}$$

$$P(X=1) = P(\text{one defective and 3 good items})$$

$$= \frac{^3C_1 \times {^7C_3}}{^{10}C_4} = \frac{3 \times 7!}{3!4!} = \frac{1}{2}$$

$$P(X=2) = P(\text{2 defective and 2 good items})$$

$$= \frac{^3C_2 \times {^7C_2}}{^{10}C_4} = \frac{3}{10}$$

$$P(X=3) = P(\text{3 defective and 1 good item})$$

$$= \frac{^3C_3 \times {^7C_1}}{^{10}C_4} = \frac{7}{^{10}C_4} = \frac{1}{30}$$

The probability distribution of Random Variable 'X' as follow

$X = x_i$	0	1	2	3
$P(X=x_i)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{30}$

(6) Let $f(x) = 3x^2$ when $0 \le x \le 1$ be the probability density function of a continuous random variable X. Determine 'a' and 'b' such that

i) $P(X \le a) = P(X > a)$ (ii) $P(X > b) = 0.05$

Sol:-

Given data $f(x) = 3x^2 \quad 0 \le x \le 1$

i) $P(X \le a) = P(X > a)$

$$\int_0^a f(x)\,dx = \int_a^1 f(x)\,dx$$

$$\int_0^a 3x^2\,dx = \int_a^1 3x^2\,dx$$

$$\left[\frac{\cancel{3}x^3}{\cancel{3}}\right]_0^a = \left[\frac{\cancel{3}x^3}{\cancel{3}}\right]_a^1$$

$$a^3 = [1 - a^3]$$

$$2a^3 = 1$$

$$a^3 = \frac{1}{2}$$

$$a = \left(\frac{1}{2}\right)^{\frac{1}{3}}$$

(11)

$$P(X>b)=0.05$$

$$\int_b^1 3x^2\,dx = 0.5$$

$$\left[\frac{3x^3}{3}\right]_b^1 = 0.05$$

$$1-b^3 = 0.05$$

$$1-0.05 = b^3$$

$$b^3 = 0.05$$

$$b = (0.05)^{1/3}$$

$\overline{P(B)}$

b) $P(A^c/B) = \frac{P(A^c \cap B)}{P(B)} = \frac{P(B) - P(A \cap B)}{P(B)}$

$= \frac{1/3 - 1/3}{4/3}$ $= \frac{5-3}{15} \cdot \frac{2}{2/3}$

8.) $P\left(\bigcap_{i=1}^{n} A_i\right) \geq \sum_{i=1}^{n} P(A_i) - (n-1)$

Methematic induction method is used to prove this
For any event, $E \in S$, $0 \leq P(E) \leq 1$

Consider the two event $(n=2)$ then
$P(A_1 \cup A_2) \leq 1$ $\because A_1 \cup A_2$ is an event

By using addition theorem
$P(A_1) + P(A_2) - P(A_1 \cap A_2) \leq 1$
$P(A_1) + P(A_2) \leq 1 + P(A_1 \cap A_2)$
we can write the above equation

$P(A_1 \cap A_2) \geq P(A_1) + P(A_2) - 1$
$\therefore$ The statement is true for $n=2$
let us assume that statement is true for $n=k$
i.e. $P\left(\bigcap_{i=1}^{k} A_i\right) \geq \sum_{i=1}^{k} P(A_i) - (k-1)$
then

then
$P\left(\bigcap_{i=1}^{k+1} A_i\right)$ $P\left(\bigcap_{i=1}^{k} (A_i \cap A_{k+1})\right) \geq$

Cosm Unit 1

Q. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

S:- Let A and B be any two events. To write $A \cup B$ as the union of three mutually exclusive events: $A \cap B^c$, $A \cap B$ and $A^c \cap B$

$$A \cup B = (A \cap B^c) \cup (A \cap B) \cup (A^c \cap B) \quad --(1)$$

By axioms 3:

$$P(A \cup B) = P(A \cap B^c) + P(A \cap B) + P(A^c \cap B) \quad --(2)$$

Now,

$$A = (A \cap B^c) \cup (A \cap B)$$

$$B = (A^c \cap B) \cup (A \cap B)$$

Therefore

$$P(A) = P(A \cap B^c) \cup P(A \cap B) \quad --(3)$$

and $$P(B) = P(A^c \cap B) \cup P(A \cap B) \quad --(4)$$

By equation 3 and 4, we get

$$P(A) + P(B) = P(A \cap B^c) + P(A^c \cap B) + 2P(A \cap B) \quad --(5)$$

Compare eq (2) and (5),

$$P(A) + P(B) = P(A \cup B) + P(A \cap B)$$

So, $$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

52 - 26 - 13 - $\frac{3}{2}$ Contd

Q. Explain Baye's theorem :-

→ Baye's theorem is a method to revise the probability of an event given additional information. Baye's theorem calculates a conditional probability called a posterior or revised probability.

- Bayes theorem is a result in probability theory that relates conditional probabilities. If A and B denote two events, $P(A|B)$ denotes the conditional probability of A occuring, given that B occurs. The two conditional probabilities $P(A|B)$ and $P(B|A)$ are in general different.

- Baye's theorem gives a relation between $P(A|B)$ and $P(B|A)$. An important application of Bayes theorem is that it gives a rule how to update or revise the strength of evidence-based belief in light of new evidence a posteriori.

- A prior probability is an initial probability value originally obtained before any additional information is obtained.
- A Posterior probability is a probability value that has been revised by using additional information that is later obtained.

Proof:-

conditional probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad -(1)$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \quad -(2)$$

from ① & ②

$$P(B/A) \cdot P(A) = P(A/B)\, P(B).$$

$$\therefore \boxed{P(A/B) = \frac{P(B/A) \cdot P(A)}{P(B)}}$$

(Q 43 > Q 44) Q 45

~~Q43~~ Two boxes ~~B_1 and~~ B_2 ~~contain 100 and 200 light bulbs respectively.~~

Q. What is random variable ? ~~what is~~ Explain discrete random variable.

Ans:- The distribution function $F(x)$ or the density $f(x)$ completely characterizes the behaviour of a random variable x. The concept of a random variable will enable us to replace the original probability space with one in which event are set of numbers.

- Whenever you run and experiment, flip a coin, roll a die, pick a card, you assign a number to represent the value to the outcome that you get. This assign is called a random variable.

- A random variable is a variable x that assigns a real number [x], for each and every outcome of a random experiment. If s is the sample space containing all the 'n' outcomes [$e_1, e_2, e_3, \ldots e_i$, ... of random experiment and x is a random variable defined as a function $X(e)$ on S, then for every outcome e_i (where $i = 1, 2, 3, \ldots, n$) that is in S the random variable $X(e_i)$ will assign a real value x_i

- Advantages of random variable is that user can define certain probability function that make it both convenient and easy to compute the probabilities of various events

→ Discrete Random variable:-

The random variable is called a discrete random variable if it is defined over a sample space having a finite or a countable infinite number of sample points. In this case, random variable takes on discrete values and it is possible to enumerate all the values it may assume.

- A discrete random variable can only have a specific (or finite) number of numerical values

- we can have infinite discrete random variable if we think about things that we know have an estimated number. Think about the number of stars in the universe. we know that there are not a specific number that we have a way to count so this is an example of an infinite discrete random variable.

- Another example would be with investments with share market. If you were to invest ₹1 lakh at the start of year, you could only estimate the amount you would have at the end of year.

166

Q. Explain difference between discrete and continu random variable

D	C
(1) It uses countable set	It uses set of interval on R
(2) F is set of all subset of Ω	F is made from sub-interv of Ω with set operations
(3) For a set $A \in F$ $P(A) = \sum_{\omega \in A} p(\omega)$	for a set $A \in F$, $P(A) = \int_{A}^{-A} f_X(x)\,dx$
(4) distribution function (cdf): $F_X(x) = \sum_{\omega \le x} p(\omega)$	(4) distribution function (cdf): $F_X(x) = \int_{-\infty}^{x} f_X(x)\,dt$

Q. 60) 66) 67) 68

By equations (3) and (4), we get

$$P(A) + P(B) = P(A \cap B^c) + P(A^c \cap B) + 2P(A \cap B) \quad \ldots (5)$$

Compare the equation (2) and (5),

$$P(A) + P(B) = P(A \cup B) + P(A \cap B)$$

So, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Q.22 For any three arbitrary events A, B, C prove that :
$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C)$. [JNTU : Nov.-09, 10]

Ans. :

$$\begin{aligned}
P(A \cup B \cup C) &= P[(A \cup B) \cup C] \\
&= P(A \cup B) + P(C) - P[(A \cup B) \cap C] \\
&= P(A) + P(B) - P(A \cap B) + P(C) \\
&\quad - P[(A \cap C) \cup (B \cap C)] \\
&= P(A) + P(B) + P(C) - P(A \cap B) \\
&\quad - P(A \cap C) - P(B \cap C) - P((A \cap C) \cap (B \cap C)) \\
&= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\
&\quad - P(B \cap C) + P(A \cap B \cap C)
\end{aligned}$$

Q.23 If $P(A) = \frac{1}{5}$, $P(B) = \frac{2}{3}$, $P(A \cap B) = \frac{1}{15}$
Find a) $P(A \cup B)$ b) $P(A^c \cap B)$ c) $P(A \cap B^c)$.

Ans. : a) $P(A \cup B)$

$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$$

$$= \frac{1}{5} + \frac{2}{3} - \frac{1}{15}$$

$$= \frac{3 + 2 \times 5}{15} - \frac{1}{15}$$

$$= \frac{13 - 1}{15} = \frac{12}{15}$$

- A posterior probability is a probability value that has been revised by using additional information that is later obtained.

- Suppose that $B_1, B_2, B_3 \ldots B_n$ partition the outcomes of an experiment and that A is another event. For any number, k, with $1 \leq k \leq n$, we have the formula :

$$P(B_k/A) = \frac{P(A/B_k) \cdot P(B_k)}{\sum_{i=1}^{n} P(A/B_i) \cdot P(B_i)}$$

Q.43 Two boxes B_1 and B_2 contain 100 and 200 light bulbs respectively. The first box (B_1) has 15 defective bulbs and the second 5. Suppose a box is selected at random and one bulb is picked out.

a) What is the probability that it is defective ?

b) Suppose we test the bulb and it is found to be defective. What is the probability that it came from box 1 ?

Ans. :

a) Probability that it is defective : Box B_1 has 85 good and 15 defective bulbs. Similarly box B2 has 195 good and 5 defective bulbs.

Let D = "Defective bulb is picked out".

Then,

$$P(D/B_1) = \frac{15}{100} = 0.15, \; P(D/B_2) = \frac{5}{200} = 0.025.$$

Since a box is selected at random, they are equally likely.

$$P(B_1) = P(B_2) = \frac{1}{2}$$

Thus B_1 and B_2 form a partition and using above equation, we obtain

$$P(D) = P(D/B_1)\,P(B_1) + P(D/B_2)\,P(B_2)$$

$$= 0.15 \times \frac{1}{2} + 0.025 \times \frac{1}{2} = 0.0875$$

Thus, there is about 9% probability that a bulb picked at random is defective.

b) Probability that it came from box 1 :

$$P(B_1/D) = \frac{P(D/B_1)\,P(B_1)}{P(D)} = \frac{0.15 \times 1/2}{0.0875} = 0.8571$$

Q.44 A mechanical factory production line is manufacturing bolts using three machines, A, B and C. The total output, machine A is responsible for 25%, machine B for 35% and machine C for the rest. The machines that 5% of the output from machine A is defective, 4% from machine B and 2% from machine C. A bolt is chosen at random from the production line and found to be defective. What is the probability that it came from

I. machine A II. machine B III. machine C ?

Ans. : Let

D = {bolt is defective},

A = {bolt is from machine A},

B = {bolt is from machine B},

C = {bolt is from machine C}.

Given data : P(A) = 0.25, P(B) = 0.35, P(C) = 0.4.

P(D|A) = 0.05, P(D|B) = 0.04, P(D|C) = 0.02.

From the Bayes' Theorem :

$$P(A/D) = \frac{P(D/A) \times P(A)}{P(D/A) \times P(A) + P(D/B) \times P(B) + P(D/C) \times P(C)}$$

$$= \frac{0.05 \times 0.25}{0.05 \times 0.25 + 0.04 \times 0.35 + 0.02 \times 0.4}$$

$$= \frac{0.0125}{0.0125 + 0.014 + 0.008}$$

P(A/D) = 0.3621

Similarly :

$$P(B/D) = \frac{P(D/B) \times P(B)}{P(D/A) \times P(A) + P(D/B) \times P(B) + P(D/C) \times P(C)}$$

$$= \frac{0.04 \times 0.35}{0.05 \times 0.25 + 0.04 \times 0.35 + 0.02 \times 0.4}$$

$$= \frac{0.014}{0.0125 + 0.014 + 0.008} = \frac{0.014}{0.0345}$$

P(B/D) = 0.4057

$$P(C/D) = \frac{P(D/C) \times P(C)}{P(D/A) \times P(A) + P(D/B) \times P(B) + P(D/C) \times P(C)}$$

$$= \frac{0.02 \times 0.4}{0.05 \times 0.25 + 0.04 \times 0.35 + 0.02 \times 0.4}$$

$$= \frac{0.008}{0.0125 + 0.014 + 0.008} = \frac{0.008}{0.0345}$$

P(C/D) = 0.2318

cal factory production line ...
ts using three machines, A, B ...
utput, machine A is responsible ...
B for 35% and machine C for ...
nes that 5% of the output from ...
tive, 4% from machine B and ...
C. A bolt is chosen at random ...
on line and found to be ...
the probability that it came ...
machine B iii. machine C ?

is defective},

is from machine A},

is from machine B},

is from machine C}.

25, $P(B) = 0.35$, $P(C) = 0.4$.

$D|B) = 0.04$, $P(D|C) = 0.02$.

em :

$$\frac{P(D/A)\times P(A)}{) + P(D/B)\times P(B) + P(D/C)\times P(C)}$$

$$\frac{.05\times 0.25}{.04\times 0.35 + 0.02\times 0.4}$$

$$+ 0.008$$

$$\frac{P(D/B)\times P(B)}{P(D/B)\times P(B) + P(D/C)\times P(C)}$$

$$\frac{0.35}{0.35 + 0.02\times 0.4}$$

$$08 = \frac{0.014}{0.0345}$$

$$\frac{D/C)\times P(C)}{D/B)\times P(B) + P(D/C)\times P(C)}$$

$$\frac{.4}{.35 + 0.02\times 0.4}$$

$$8 = \frac{0.008}{0.0345}$$

Q.45 At a certain university, 4% of men are over 6 feet tall and 1% of women are over 6 feet tall. The total student population is divided in the ratio 3:2 in favour of women. If a student is selected at random from among all those over six feet tall, what is the probability that the student is a woman ?

Ans. : Let us assume following :

M = {Student is Male},

F = {Student is Female},

T = {Student is over 6 feet tall}.

Given data :

$P(M) = 2/5$,

$P(F) = 3/5$,

$P(T|M) = 4/100$

$P(T|F) = 1/100$.

We require to find $P(F|T)$?

Using Bayes' Theorem we have :

$$P(F/T) = \frac{P(T/F)\,P(F)}{P(T/F)\,P(F) + P(T/M)\,P(M)}$$

$$= \frac{\frac{1}{100}\times\frac{3}{5}}{\frac{1}{100}\times\frac{3}{5} + \frac{4}{100}\times\frac{2}{5}} = \frac{\frac{3}{500}}{\frac{3}{500} + \frac{8}{500}}$$

$$P(F/T) = \frac{3}{11}$$

Q.46 A pair of dice is rolled. If the sum of 9 has appeared, find the proabability that one of the dice shows 3.

Ans. : Let A = The event that the sum is 9

B = The event the one of dice shows 3.

Exhaustive cases $= 6^2 = 36$.

Favorable cases of the event A = (3, 6), (6, 3), (4, 5), (5, 4).

So $P(A) = 4.36$

$$P(A) = \frac{1}{9}$$

Favorable case for the event $A \cap B$ = (3, 6), (6, 3)

Hence $P(A \cap B) = \frac{2}{36} = \frac{1}{18}$

But $P(A \cap B) = P(A) \times P(B/A)$

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

$$P(B/A) = \frac{1/18}{1/9} = \frac{1}{18}\times\frac{9}{1}$$

$$P(B/A) = 1/2$$

Q.47 If A_1 and A_2 are equally lik... exclusive and exhaustive events and P... $P(B/A_2) = 0.3$, find $P(A_1/B)$.

Ans. : Since A_1 and A_2 are equally ... $P(A_1) = P(A_2)$. Further, they are mutu...

$$P(A_1) + P(A_2) = 1$$

$$P(A_1) = P(A_2) = 0.5$$

$$= \frac{0.5\times 0.2}{0.5\times 0.2 + 0.5\times ...}$$

$$= \frac{0.1}{0.1 + 0.15} = \frac{0.1}{0.25}$$

$$P(A_1/B) = 0.4$$

Q.48 10% of the bulbs produced ... of red colour and 2% are red and ... bulb is picked up at random, ... probability of its being defective if ...

Ans. :

Let A and B be the events that the b... detective, respectively.

$$P(A) = 10/100 = 1/10$$

$$P(A \cap B) = 2/100 = 1/50$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1/50}{1/10}$$

Thus the probability of the pick... being defeftive, if it is red, is 1/5.

1.6 : Random Variables and Distributions : Concep... Variable

Q.49 Define discrete sample spa...

Ans. : If a sample space contain... possibilities or an unending seq... elements as there are whole nu... discrete sample space.

om variable.

a continuous rando
in an interval

distribution for th

head side up.

il side up.

coin to land hea

e number of
. Determine

$$= \frac{1}{16}+\frac{4}{16}+\frac{6}{16}+\frac{4}{16}$$

$$= \frac{1+4+6+4}{16} = \frac{15}{16}$$

$$F(4) = f(0) + f(1) + f(2) + f(3) + f(4)$$

$$= \frac{1}{16}+\frac{4}{16}+\frac{6}{16}+\frac{4}{16}+\frac{1}{16}$$

$$= \frac{1+4+6+4+1}{16} = \frac{16}{16} = 1$$

Q.55 A shipment of 8 similar microcomputers to a. retail outlet contains 3 that are defective. If a school makes a random purchase of 2 of these computers, find the probability distribution for the number of defectives.

Ans. : Let X be a random variable whose values x are the possible numbers of defective computers purchased by the school. Then x can be any of the numbers 0, 1, and 2. Now

Number of ways of choosing any 2 :

$$^8C_2 = \frac{8!}{((8-2)!2!)} = 28$$

(This will be the denominator)

$$P[0] = {}^3C_0 \times {}^5C_2/28$$

$$= (3!/((3-0)!0!) \times (5! \times ((5-2)!2)))/28$$

$$= 5/14$$

$$P[1] = {}^3C_1 \times {}^5C_{1/28} = 15/28$$

$$P[2] = {}^3C_2 \times {}^5C_{0/28} = 3/28$$

Q.56 If a random variable X takes the values 1, 2, 3 and 4 such that 2P (X = 1) = 3P (X = 2) = P (X = 3) = 5P (X = 4), derive the probability distribution function of X.

[JNTU : March-17, Marks 5]

Assume P (X = 3) = α By the given equation

$$1) = \frac{\alpha}{2} \quad P(X = 2) = \frac{\alpha}{3} \quad P(X = 4) = \frac{\alpha}{5}$$

a probability distribution (and mass function)

$$\sum P(x) = 1$$

$$P(1) + P \quad) + P(4) = 1$$

$$\frac{\alpha}{2} + \quad 1 \quad \Rightarrow \frac{61}{30}\alpha = 1 \quad \Rightarrow \alpha = \frac{30}{61}$$

$$P(X = 1) = \frac{15}{61};\ P(X = 2) = \frac{10}{61};\ P(X$$

$$P(X = 4) = \frac{6}{61}$$

- The probability distribution giv

X	1	2	
P(X)	15/61	10/61	30

1.7 : Continuous Probability Statistical Independence

Q.57 What is Continuous probabil

Ans. : Continuous probability dist of distribution that deals with co data or random variables. The co variables deal with different kinds o

Q.58 What is a discrete probabil What two conditions determine distribution?

Ans. : • Discrete distribution is distribution where sample space is random variable in such a case can values. For example: number of peop coins etc.

- Discrete Probability distribution values a random variable can a corresponding probabilities of the va
- The two conditions are

1. Sum of all probabilities should be
2. Each probability must be greater to zero.

Q.59 Explain difference between continuous random variable.

Ans. :

Sr. No.	Discrete	C
1.	It uses countable set	It uses on R.

2.	F is set of all subset of Ω.	F is made from sub-intervals of Ω with set operations.
3.	For a set $A \in F$, $P(A) = \sum_{\omega \in A} p(\omega)$	For a set $A \in F$, $p(A) = \int_A f_X(x)\,dx$
4.	Distribution function (Cdf) : $F_X(x) = \sum_{\omega \le x} p\omega$	Distrinbution function (Cdf) : $F_X(x) = \int_{-\infty}^{x} f_X(x)\,dt$

Q.60 Explain statistical independence

Ans. : Two variates A and B are statistically independent if the probability $P(A|B)$ of A given B satisfies $P(A|B) = P(A)$

in which case the probability of A and B is just

$$P(AB) = P(A \cap B) = P(A)\,P(B)$$

- Statistical independence means one event conveys no information about the other; statistical dependence means there is some information.
- Statistically independent is not the same as mutually exclusive: if A and B are mutually exclusive, then they can't be independent, unless one of them is probability 0 to start with :

$P(r)(()A \cap B) = 0 = \Pr(()A)\Pr(()B)$

Q.61 Three cards are drawn in succession from a deck without replacement. Find the probability distribution for the number of spades.

Ans. : X = number of spades in the three draws.

Let S and N stand for a spade and not a spade, respectively.

$$P(X = 0) = P(NNN)$$
$$= (39/52)\,(38/51)\,(37/50)$$
$$= 703/1700$$

$$P(X = 1) = P(SNN) + P(NSN) + P(NNS)$$
$$= 3(13/52)(39/51)(38/50)$$
$$= 741/1700$$

$$P(X = 3) = P(SSS)$$
$$= (13/52)(12/51)(11/50)$$
$$= 11/850$$

$$P(X = 2) = 1 - \frac{703}{1700} - \frac{741}{1700} - \frac{11}{850} =$$

Q.62 From a box containing 4 black ... green balls, 3 balls are drawn in suc... ball being replaced in the box be... draw is made. Find the probability d... the number of green balls.

Ans. : • Let X denotes the number of ... the three draws.

- Let G and B stand for the colors ... black, respectively.

Sample Event	x	
BBB	0	(2/3)3
GBB	1	(1/3)
BGB	1	(1/3)
BBG	1	(1/3)
BGG	2	(1/3)2
GBG	2	(1/3)2
GGB	2	(1/3)2
GGG	3	(1/3)3

Q.63 A traffic engineer is interested i... of vehicles reaching a particular cros... periods of relatively low traffic flow. ... finds that the number of vehicles X ... crossroads per minute is gover... probability distribution:

x	0	1	2
P(X = x)	0.37	0.39	0.19

Calculate the expected value, the ... the standard deviation of the rand...

Ans. : The expected value, the va... standard deviation of the random vaia...

x	x^2	P(X = x)
0	0	0.37
1	1	0.39
2	4	0.04
3	9	0.09
4	16	0.01

$\times 0.19+3\times 0.04+4\times 0.01$

$xP(X=x)\Big]^2$

$0.19+9\times 0.04+16$

ven by $\sigma = \sqrt{V(X)} =$

gram of the following ch represents the s from a video store

... be a random variable with the following probability distribution

X	−3	6	9
P(X = x)	1/6	1/2	1/3

then evaluate $E(2X+1)^2$

Ans. : We know that :

$$E(2X+1)^2 = E(4X^2+4X+1) = E4(X^2)+E4(X)+1$$

$$E(X) = (-3)\frac{1}{6}+(6)\frac{1}{2}+(9)\frac{1}{3} = \frac{11}{12}$$

$$E(X^2) = (-3)^2\frac{1}{6}+(6^2)\frac{1}{2}+(9^2)\frac{1}{3} = \frac{93}{2}$$

$$E(2X+1)^2 = (4)\times(93/2)+(4)(11/2)+1 = 209$$

Q.66 A class contains of 6 girls and 10 boys. If a committee of 3 is chosen at random from the class, find the probability that (i) 3 boys are selected, (ii) exactly 2 girls are selected.

Ans. : Total number of student (n) = 16

Then n(s) = number of ways of choosing 3 from 16

$$= {}^{16}C_3$$

i) 3 boys are selected : We can write ${}^{10}C_3$ here.

$$n(E) = {}^{10}C_3 \text{ Therefore}$$

$$P(E) = \frac{{}^{10}C_3}{{}^{16}C_3} = \frac{10\times 9\times 8}{16\times 15\times 14} = \frac{3}{14}$$

ii) exactly 2 girls are selected

Then $n(E) = {}^{6}C_2\times{}^{10}C_1$

$$P(E) = \frac{{}^{6}C_2\times{}^{10}C_1}{{}^{16}C_3} = \frac{15}{56}$$

Q.67 Three groups of students contain respectively 3 girls and 1 boy; 2 girls and 2 boys; 1 girl and 3 boys; one student is selected at random from each group. Find the probability of selecting 1 girl and 2 boys.

Ans. : If G denotes selection of girls and B denotes selection of a boy then the various cases of selection can be GBB respectively from the three groups, or BBG respectively from the three groups.

So, P(GBB) = P(G from group 1) × P(B from group II) × P(B from group III)

$$P(GBB) = \frac{^3C_1}{^4C_1} \times \frac{^2C_1}{^4C_1} \times \frac{^3C_1}{^4C_1} = \frac{3}{2} \times \frac{2}{4} \times \frac{3}{4} = \frac{9}{32}$$

Similarly,

$$P(BGB) = \frac{1}{4} \times \frac{2}{4} \times \frac{3}{4} = \frac{3}{32}$$

$$P(BBG) = \frac{1}{4} \times \frac{2}{4} \times \frac{1}{4} = \frac{1}{32}$$

Required probability = P(GBB) + P(BGB) + P(BBG)

$$= \frac{9}{32} + \frac{3}{32} + \frac{1}{32} = \frac{13}{32}$$

Q.68 A batch of 100 manufactured components is checked by an inspector who examines 10 components selected at random. If none of the 10 component's is defective, the inspector accepts the whole batch. Otherwise, the batch is subjected to further inspection. What is the probability that a batch containing 10 defective components will be accepted ?

Ans. : • Let the N denote the number of ways of indiscriminately selecting 10 components from a batch of 100 components. Then N is given by

$$N = C(100, 10)$$

$$= \frac{100!}{(100-10)! \times 10!} = \frac{100!}{90! \times 10!}$$

- Let E denote the event "the batch containing 10 defective components is accepted by the inspector".
- The number of ways that E can occur is the number of ways of selecting 10 components from the 90 non-defective components and no components from the 10 defective component's.
- This number, N(E) is given by

$$N(E) = C(90, 10) \times C(10, 0) = C(90, 10)$$

$$= \frac{90!}{(90-10)! \times 10!} = \frac{90!}{80! \times 10!}$$

The probability of event E is given by :

$$P(E) = \frac{N(E)}{N} = \frac{90!}{80! \times 90!} \times \frac{90! \times 10!}{100!} = \frac{90! \times 90!}{100! \times 80!}$$

$$P(E) = 0.3305$$

Q.69 Two coins are tossed. Let A d "at most one head on the two tos denote the event "one head and o tosses". Are A and B independent e

Ans. : The sample space of the expe HT, TH, TT}

Now events are defined as follows :

A = {HT, TH, TT} B =

and A ∩ B = { HT, TH}

Thus :

$$P(A) = \frac{3}{4} \qquad P(B) =$$

$$P(A \cap B) = \frac{2}{4} = \frac{1}{2}$$

$$P(A)\,P(B) = \frac{8}{3}$$

Since $P(A \cap B)$ $P(A)\,P(B)$, we concl and B are not independent.

Q.70 A die is thrown 3 times. I considered as success find the atleast 2 success.

Ans. :

$$p = \frac{1}{6},\ q = \frac{5}{6},\ n = 3$$

P (at least 2 success) = $P(X \geq 2)$ = P

$$= 3C_2 \left(\frac{1}{6}\right)^2 \frac{5}{6} + 3C_3 ($$

Fill in the Blanks for Mid

Q.1 A ______ variable takes which is determined by t random experiment.

Q.2 A pictorial representation manipulations with events using ______ diagrams.

Q.3 Statistical data, generated in be very useful for studyin the distribution if presente tabular and graphic displa plot

UNIT-II

Mathematical Expectation &

Discrete probability Distributions

Short Answer questions

(1) Define Expectation & variance of random variable

Expectation for Discrete Variable

A Random variable 'X' assumes the values $x_1, x_2, x_3 --- x_n$ with respective probabilities $p_1, p_2 -- p_n$ Then $E(X) = x_1p_1 + x_2p_2 + --- + x_np_n$

$$E(X) = \sum_{i=1}^{n} x_i p_i$$

Continuous Case:

'x' is Continuous Random variable and $f(x)$ is probability density function Then

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

(2) P.T $E(aX+b) = aE(X)+b$ and

$Var(X+k) = Var(X)$

Sol:-

$$E(aX+b) = \sum_{i=1}^{n} (ax_i+b)\, P(X=x_i)$$

$$= a\sum_{i=1}^{n} x_i P(X=x_i) + b\sum_{i=1}^{n} P(X=x_i)$$

$$= a\; E(X) + b(1)$$

$$= a\; E(X) + b$$

$$Var(X+k) = \int_{-\infty}^{\infty} (x+k)^2 f(x)\,dx - \left[\int_{-\infty}^{\infty} (x+k) f(x)\,dx\right]^2$$

$$= \int_{-\infty}^{\infty} (x^2+2kx+k^2) f(x)\,dx - \left[\int_{-\infty}^{\infty} x f(x) + k \int_{-\infty}^{\infty} f(x)\,dx\right]^2$$

$$= \int_{-\infty}^{\infty} x^2 f(x)\,dx + 2k \int_{-\infty}^{\infty} x f(x)\,dx + k^2 - \left[\int_{-\infty}^{\infty} x f(x)\,dx + k\right]^2$$

$$= E(X^2) + 2k\,E(X) + k^2 - [E(X)+k]^2$$

$$= E(X^2) + \cancel{2k\,E(X)} + \cancel{k^2} - [E(X)]^2 - \cancel{k^2} - \cancel{2k\,E(X)}$$

$$= E(X^2) - E(X)^2$$

$$Var(X+k) = Var(X)$$

(3) Define Binomial distribution & Poisson distribution

Def: A Random Variable X has a Binomial distribution if it assumes only non-negative values and its probability density function is given by

$$P(X=r) = P(r) = \begin{cases} {}^{n}C_r\, p^r q^{n-r} & r = 0,1,2 \ldots n, \quad q = 1-p \\ 0 & \text{otherwise} \end{cases}$$

Poisson distribution:

A random variable 'X' is said to follow a poisson distribution if it assumes only non-negative values and its probability density function is given by

$$p(x,\lambda) = p(X=x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^{x}}{x!} & x=0,1,2,3 \\ 0 & \text{otherwise} \end{cases}$$

Here $\lambda>0$ is called the parameter of the distribution.

(4) The mean of Binomial distribution is '3' & var $\frac{9}{4}$ find 'n'

mean $np=3$ var $npq=\frac{9}{4}$

$$\frac{\cancel{n}pq}{\cancel{n}p} = \frac{\frac{\cancel{9}^{3}}{4}}{\cancel{3}} = \frac{3}{4} \qquad q=\frac{3}{4} \quad p=\frac{1}{4}$$

$np=3$

$n\frac{1}{4}=3 \qquad n=12$

(5) The prob of poisson variate 'X' taking the values 1 & 2 are equal Find i) M (ii) $P(X\geq 1$

(iii) $P(1<X<4)$

Sol:

$$P(X=1) = P(X=2)$$

$$\frac{\cancel{e^{-\lambda}} \cancel{\lambda^1}}{1!} = \frac{\cancel{e^{-\lambda}} \lambda^{\cancel{2}}}{2!} \Rightarrow \lambda = 2$$

$$\text{mean} = \mu = \lambda = 2$$

(ii) $P(X \geq 1) = 1 - P(X=0)$

$$= 1 - \frac{e^{-2} 2^0}{0!} = 1 - \frac{1}{e^2}$$

(iii) $P(1 < X < 4) = P(X=2) + P(X=3)$

$$= \frac{e^{-2} 2^{\cancel{2}}}{\cancel{2}!} + \frac{e^{-2} 2^3}{3!}$$

$$= \frac{2}{e^2} + \frac{e^{-2}\, 2 \times \cancel{2} \times 2}{3 \times \cancel{2}} = \frac{2}{e^2} + \frac{4}{3e^2}$$

6 Derive mean of Geometric distribution

Geometric distribution

$$P(X=x) = q^{x-1} p \quad \text{where } q = 1-p$$

where
- p is success of out come
- q is failure of out come
- x is number of trails required to get a first success

Mean:

$$\mu = E(X) = \sum_{x=1}^{\infty} x\, P(X=x)$$

$$= \sum_{x=1}^{\infty} x q^{x-1} p$$

$$= p[1 \cdot q^{1-1} + 2q^{2-1} + 3q^{3-1} + \cdots]$$

$$= p[1 + 2q + 3q^2 + 4q^3 + \cdots]$$

$$= p[(1-q)^{-2}]$$

$$= \frac{p}{p^2} = \frac{1}{p}$$

$$\text{mean} = \mu = \frac{1}{p}$$

long Answer questions

1. A Random variable 'x' has following probability function

x	0	1	2	3	4	5	6	7
$p(x)$	0	K	$2K$	$2K$	$3K$	K^2	$2K^2$	$7K^2+K$

i) Determine K (ii) Evaluate $P(X<6)$, $P(X\geq 6)$, $P(0<X<5)$

(iii) if $P(X\leq \cancel{K})>\frac{1}{2}$ find the minimum value of 'k'

iv) Determine the distribution function of 'X'

v) mean vi) Variance

Sol:-

i) $\sum_{x=1}^{7} p(x)=1$

$0+K+2K+2K+3K+K^2+2K^2+7K^2+K=1$

$10K^2+9K=1$

$10K^2+9K-1=0$

$(10K-1)(K+1)=0$

$K=\frac{1}{10}=0.1$ (Since $p(x)\geq 0$

so $K\neq -1$)

So $K=0.1=\frac{1}{10}$ sub in above distribution

table

x	0	1	2	3	4	5	6	7
$P(x)$	0	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{1}{100}$	$\frac{2}{100}$	$\frac{17}{100}$

(ii) $P(X<6) = P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5)$

$$= 0 + \frac{1}{10} + \frac{2}{10} + \frac{2}{10} + \frac{3}{10} + \frac{1}{100} = 0.81$$

$P(X \geq 6) = 1 - P(X<6) = 1 - 0.81 = 0.19$

$P(0<X<5) = P(X=1) + P(X=2) + P(X=3) + P(X=4)$

$$= \frac{1}{10} + \frac{2}{10} + \frac{2}{10} + \frac{3}{10} = \frac{8}{10} = 0.8$$

(iii) The required minimum value of 'k'

$P(X \leq 1) = P(X=0) + P(X=1) = 0 + \frac{1}{10} = 0.1$

$P(X \leq 2) = P(X=0) + P(X=1) + P(X=2) = 0.1 + \frac{2}{10} = 0.3$

$P(X \leq 3) = P(X \leq 2) + P(X=3) = 0.3 + \frac{2}{10} = 0.5$

$P(X \leq 4) = P(X \leq 3) + P(X=4) = 0.5 + \frac{3}{10} = 0.8 > 0.5 = \frac{1}{2}$

$P(X \leq 4) = 0.8 > 0.5 = \frac{1}{2}$

minimum value of $k = 4$

iv) The distribution function of X is given by following table

X	$F(x) = P(X \le x)$
0	0
1	$0 + \frac{1}{10} = \frac{1}{10}$
2	$\frac{1}{10} + \frac{2}{10} = \frac{3}{10}$
3	$\frac{3}{10} + \frac{2}{10} = \frac{5}{10}$
4	$\frac{5}{10} + \frac{3}{10} = \frac{8}{10}$
5	$\frac{8}{10} + \frac{1}{100} = \frac{81}{100}$
6	$\frac{81}{100} + \frac{2}{100} = \frac{83}{100}$
7	$\frac{83}{100} + \frac{17}{100} = \frac{100}{100} = 1$

v) Mean $= \sum_{i=0}^{7} x_i p_i$

$= 0 \times 0 + 1 \times \frac{1}{10} + 2 \times \frac{2}{10} + 3 \times \frac{2}{10} + 4 \times \frac{3}{10} + 5 \times \frac{1}{100} + 6 \times \frac{2}{100} + 7 \times \frac{17}{100}$

$= 3.66$

vi) $Var(X) = E(X^2) - [E(X)]^2$

$E(X^2) = \sum_{i=0}^{7} x_i^2 p_i$

$= 0^2 \times 0 + 1^2 \times \frac{1}{10} + 2^2 \times \frac{2}{10} + 3^2 \times \frac{2}{10} + 4^2 \times \frac{3}{10} + 5^2 \times \frac{1}{100} + 6^2 \times \frac{2}{100} + 7^2 \times \frac{17}{100} = 16.8$

$Var(X) = 16.8 - (3.66)^2$

$= 16.8 - 13.3956 = 3.4044$

(2) Derive mean and variance of Binomial distribution

Sol:- The Binomial probability distribution is given by

$$p(r) = n_{C_r} p^r q^{n-r} \quad r = 0, 1, 2 \ldots n \quad q = 1-p$$

$$\text{Mean of } X = \mu = E(X) = \sum_{r=0}^{n} r\, p(X=r)$$

$$= \sum_{r=0}^{n} r\, n_{C_r} p^r q^{n-r}$$

$$= 0 \times q^n + 1 \times n_{C_1} p q^{n-1} + 2 n_{C_2} p^2 q^{n-2} + \cdots + n p^n$$

$$= n p q^{n-1} + 2 \frac{n(n-1)}{2!} p^2 q^{n-2} + \cdots + n p^n$$

$$= np[q^{n-1} + (n-1) p q^{n-2} + \cdots - + p^{n-1}]$$

$$= np (q+p)^{n-1} \qquad [q+p=1]$$

$$= np$$

variance of Binomial distribution

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

$$E(X^2) = \sum_{r=0}^{n} r^2\, p(X=r)$$

$$= \sum_{r=0}^{n} r^2\, n_{C_r} p^r q^{n-r}$$

$$= \sum_{r=0}^{n} [r(r-1) + r]\, n_{C_r} p^r q^{n-r}$$

$$= \sum_{r=0}^{n} r(r-1)\, {}^{n}C_{r}\, p^{r} q^{n-r} + \sum_{r=0}^{n} r\, {}^{n}C_{r}\, p^{r} q^{n-r}$$

$$= \Big[0(0-1)\, {}^{n}C_{0}\, p^{0} q^{n-0} + 1(1-1)\, {}^{n}C_{1}\, p^{1} q^{n-1} + 2(2-1)\, {}^{n}C_{2}\, p^{2} q^{n-2} + \cdots$$

$$\cdots + n(n-1)\, {}^{n}C_{n}\, p^{n} q^{n-n}\Big] + np$$

$$= 2\, \frac{n(n-1)}{\cancel{2!}}\, p^{2} q^{n-2} + \cdots \quad + n(n-1)p^{n}\Big] + np$$

$$= n(n-1)p^{2}\Big[\, q^{n-2} + \cdots \quad + p^{n-2}\Big] + np$$

$$= n(n-1)p^{2}\,[q+p]^{n-2} + np$$

$$E(x^{2}) = n(n-1)p^{2} + np$$

$$var(x) = E(x^{2}) - (E(x))^{2}$$

$$= n(n-1)p^{2} + np - (np)^{2}$$

$$= \cancel{n^{2}p^{2}} - np^{2} + np - \cancel{n^{2}p^{2}}$$

$$= np(1-p)$$

$$var(x) = npq$$

(3) Ten Coins are tossed. Find the probability of getting greater than (or) equal to 6 heads

Sol:-

Binomial distribution

$P(X=r) = {}^{n}C_{r}\, p^{r} q^{n-r}$

prob of getting head $p = \frac{1}{2}$ $\quad q = 1-p = \frac{1}{2}$

$n = 10$

find

$P(X \geq 6) = P(X=6) + P(X=7) + P(X=8) + P(X=9) + P(X=10)$

$= {}^{10}C_6 \left(\frac{1}{2}\right)^6 \left(\frac{1}{2}\right)^{10-6} + {}^{10}C_7 \left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^{10-7}$

$+ {}^{10}C_8 \left(\frac{1}{2}\right)^8 \left(\frac{1}{2}\right)^{10-8} + {}^{10}C_9 \left(\frac{1}{2}\right)^9 \left(\frac{1}{2}\right)^{10-9}$

$+ {}^{10}C_{10} \left(\frac{1}{2}\right)^{10} \left(\frac{1}{2}\right)^{10-10}$

$= (210 + 120 + 45 + 10 + 1) \left(\frac{1}{2}\right)^{10}$

$= 0.37$

4) Derive mean and variance of poisson distribution.

Sol:-

The poisson probability distribution is given by

$$P(X=r) = \frac{e^{-\lambda}\lambda^r}{r!} \qquad r = 0,1,2 \ldots$$

$$\text{Mean} = E(X) = \mu = \sum_{r=0}^{\infty} r \frac{e^{-\lambda}\lambda^r}{r!}$$

$$= \sum_{r=0}^{\infty} \cancel{r} \frac{e^{-\lambda}\lambda^r}{\cancel{r}(r-1)!}$$

$$= e^{-\lambda} \sum_{r=1}^{\infty} \frac{\lambda^r}{(r-1)!}$$

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

$$= \lambda e^{-\lambda}\left[1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \ldots\right]$$

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

$$\text{variance}(X) = E(X^2) - [E(X)]^2$$

$$E(X^2) = \sum_{r=0}^{\infty} r^2 P(X=r)$$

$$= \sum_{r=0}^{\infty} \left(r(r-1) + r\right) P(X=r)$$

$$= \sum_{r=0}^{\infty} \left(r(r-1)\right) P(X=r) + \sum_{r=0}^{\infty} r\, P(X=r)$$

$$= \sum_{r=0}^{\infty} r(r-1)\frac{e^{-\lambda}\lambda^r}{r!} + \lambda$$

$$= e^{-\lambda}\sum_{r=0}^{\infty} \cancel{r(r-1)}\frac{\lambda^r}{\cancel{r(r-1)}(r-2)!} + \lambda$$

$$= e^{-\lambda}\sum_{r=2}^{\infty} \frac{\lambda^r}{(r-2)!} + \lambda$$

$$= e^{-\lambda}\left[\frac{\lambda^2}{0!} + \frac{\lambda^3}{1!} + \frac{\lambda^4}{2!} + \cdots\right] + \lambda$$

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

$$= \cancel{e^{-\lambda}}\lambda^2 \cancel{e^{\lambda}} + \lambda = \lambda^2 + \lambda$$

$$Var(X) = E(X^2) - [E(X)]^2$$

$$= \cancel{\lambda^2} + \lambda - \cancel{[\lambda]^2}$$

$$var(X) = \lambda$$

5) P.T poisson distribution is the limiting case of Binomial distribution

Proof:- The poisson distribution can be derived as a limiting case of the Binomial distribution under the following conditions

i) p, the probability of the occurrence of the is very small

ii) n is very very large, where 'n' is no. of trials i.e $n \to \infty$

iii) np is a finite quantity say $np = \lambda$ then λ is called the parameter of the poisson distribution.

Binomial distribution

$$P(X=r) = nc_r p^r q^{n-r}$$

$$= nc_r p^r (1-p)^{n-r}$$

$$= \frac{n(n-1)(n-2)\cdots(n-r+1)}{r!} p^r \frac{(1-p)^n}{(1-p)^r} \rightarrow (1)$$

put $np = \lambda$ then $n = \frac{\lambda}{p}$

$$p(r) = \frac{\frac{\lambda}{p}\left(\frac{\lambda}{p}-1\right)\left(\frac{\lambda}{p}-2\right)\cdots\left(\frac{\lambda}{p}-r+1\right)}{r!} \cdot p^r \cdot \frac{(1-p)^n}{(1-p)^r}$$

$$= \frac{\lambda(\lambda-p)(\lambda-2p)\cdots[\lambda-(r-1)p]}{r!\, p^r} p^r \frac{(1-p)^n}{(1-p)^r}$$

$$= \frac{\lambda(\lambda-p)(\lambda-2p)\cdots[\lambda-(r-1)p]}{r!} \frac{\left(1-\frac{\lambda}{n}\right)^n}{(1-p)^r}$$

As $n \to \infty$, $p \to 0$ so that $np = \lambda$ we have

$$p(r) = \frac{\lambda \cdot \lambda \cdots r \text{ factors}}{r!} \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n \lim_{p\to 0} \frac{1}{(1-p)^r}$$

$$= \frac{\lambda^r}{r!} \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n \quad \left[\because \lim_{p\to 0}(1-p)^r = 1 \text{ for a given } r\right]$$

$$= \frac{\lambda^r}{r!} e^{-\lambda} \quad \left[\because \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = \lim_{n\to\infty}\left[\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda} = e^{-\lambda}\right]$$

$p(r)$ = probability of 'r' successes $= \dfrac{e^{-\lambda}\lambda^r}{r!}$

This is known as poisson distribution

where $r = 0, 1, 2 \ldots$

(6) Average number of accidents on any day on a national highway is 1·6. Determine the probability that the no of accidents is

i) At least one (ii) at most one

Sol:- Given data avg = mean = $\lambda = 1.6$

poisson distribution $P(X=r) = \dfrac{e^{-\lambda}\lambda^r}{r!}$

ii) p(at ~~least~~ most one) $= P(X \le 1) = P(X=0) + P(X=1)$

$$= \frac{e^{-1.6}(1.6)^0}{0!} + \frac{e^{-1.6}(1.6)^1}{1!}$$

$$= \frac{1}{e^{1.6}}(1+1.6) = \frac{2.6}{e^{1.6}} = 0.52$$

(bi) $p(\text{at least one}) = p(X\geq 1)$

$$= 1 - p(X=0)$$

$$= 1 - \frac{e^{-1.6}(1.6)^0}{0!}$$

$$= 1 - \frac{1}{e^{1.6}} = 0.80$$

7). Six Cards are drawn from a pack of 52 Cards Find the probability that

i) at least three are diamonds

(ii) 4 are diamonds

Sol:- $P(\text{getting a diamond Card}) = \frac{13_{C_1}}{52_{C_1}} = \frac{13}{52} = \frac{1}{4}$

$$q = 1-p = 1-\frac{1}{4} = \frac{3}{4}$$

i) $p(\text{at least three are diamond}) = p(X\geq 3)$

$n=6 \quad p=\frac{1}{4} \quad q=\frac{3}{4}$

$$= 1 - p(X<3)$$

$$= 1 - [p(X=0) + p(X=1) + p(X=2)]$$

$$= 1 - \left[6_{C_0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^6 + 6_{C_1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^5 + 6_{C_2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^4\right]$$

$$= 1 - [0.177 + 0.355 + 0.2966]$$

$$= 0.1714$$

(ii) $P(4 \text{ are diamonds}) = P(X=4)$

$$= {}^6C_4\left(\frac{1}{4}\right)^4\left(\frac{3}{4}\right)^2$$

$$= 15 \times 0.0039 \times 0.56$$

$$= 0.0329$$

8) A sample of '5' items is selected at random from a box containing 15 items of which '8' are defective find i) mean (ii) variance of defective items.

$$P(\text{defective item}) = \frac{8}{15}$$

$$q = 1 - \frac{8}{15} = \frac{7}{15}$$

sample size $n = 5$

i) $\text{mean} = np = \cancel{5} \times \frac{8}{\cancel{15}_3} = \frac{8}{3} = 2.66$

(ii) $\text{variance} = npq = \cancel{5} \times \frac{8}{\cancel{15}_3} \times \frac{7}{15} = \frac{56}{45}$

9) Chebyshev's theorem:

Chebyshev's theorem states that the proportion (or) percentage of any data set that lies within 'k' standard deviation of the mean where 'k' is any positive integer >1 is atleast $1 - \frac{1}{k^2}$

Continuous probability Distribution & Fundamental Sampling Distributions

Unit-3

Short Answer questions

(1) Define normal distribution

Ans:- A random Variable X is said to have a normal distribution, if its density function is given by

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}; \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$$

Where μ = mean, σ = S.D

(2) Define Gamma & Exponential distribution

Ans:- Let X be a Continuous random variable, assuming only non-negative values, distributed according to Gamma probability density function given by

$$f(x) = \begin{cases} \frac{\beta}{\Gamma(\alpha)} (\beta x)^{\alpha-1} e^{-\beta x}, & 0 < x < \infty, \ \alpha > 0, \beta > 0 \\ 0 & , \text{otherwise} \end{cases}$$

Exponential:- A continuous random Variable X having the range $0 < x < \infty$ is said to have exponential distribution if it has a probability density is given by

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

(3) A random Sample of size 80 is taken from a Population whose S.D is 15. Find the Standard error of means.

Sol:- $n = 80, \quad \sigma = 15$

Standard error of mean $= \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{80}} = 1.68.$

(4) Find the Value of the finite Population Correction factor for $n=10$ & $N=100$

Sol:- $N = 1000, \; n = 10$

Correction factor $= \frac{N-n}{N-1} = \frac{1000-10}{1000-1} = 0.991$

(5) Define Simple, random, Purposive Sample.

Simple Sample:- Simple Sample is defind as every item in the population has equal Chance.

Random Sample:- Random Sampling is Sample that it has an equal probability of being Choosen.

Purposive Sampling:- It is a non probability sampling Method, the elements selected for the sample are Choosen by the judgment of the researcher.

long Answer question

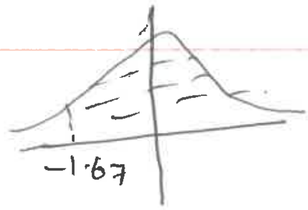
① If x is normally distributed with Mean 1 and S.D 0.6 obtain $P(x>0)$ & $P(-1.8 \le X \le 2.0)$

Sol:

Given that $\mu = 1$, $\sigma = 0.6$

① When $x = 0$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{0-1}{0.6} = -1.67$

$$\therefore P(x>0) = P(z > -1.67)$$



$$= 0.5 + A(1.67)$$
$$= 0.5 + 0.4525$$
$$= 0.9525$$

② $P(-1.8 < X \le 2.0)$

When $x = -1.8$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{-1.8-1}{0.6} = -4.67$

$x = 2.0$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{2.0-1}{0.6} = 1.67$

$$\therefore P(-1.8 < X \le 2) = P(-4.7 < z < 1.7)$$
$$= A(1.67) + A(-4.67)$$
$$= 0.4525 + 0.0691$$
$$= 0.5216$$

(2) The marks obtained by 500 students is normally distributed with mean 65% & S.D 8%. Determine how many get more than 80%.

Sol:- $\mu = 0.65, \quad \sigma = 0.08$

$P(X > 0.8)$

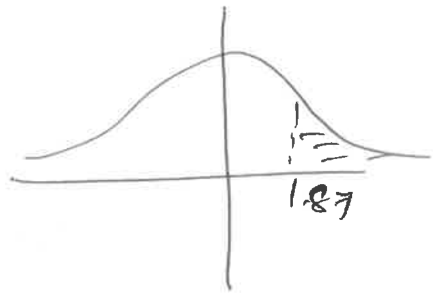
when $x = 0.8$, $z = \frac{x-\mu}{\sigma} = \frac{0.8-0.65}{0.08} = 1.87$

$\therefore P(X > 0.8) = P(z > 1.87)$

$= 0.5 - A(1.87)$

$= 0.5 - 0.4693$

$= 0.0307$



(3) Given that the mean heights of students in a class is 158 Cms with a S.D of 20cms. Find how many students height lie between 150 & 170 Cms.

Sol:- $\mu = 158, \quad \sigma = 20$

when $x = 150$, $z = \frac{x-\mu}{\sigma} = \frac{150-158}{20} = -0.4$

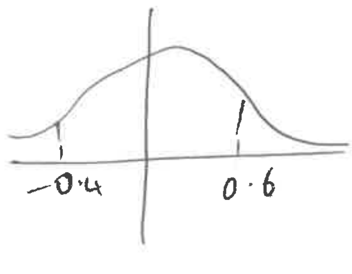
$x = 170$, $z = \frac{x-\mu}{\sigma} = \frac{170-158}{20} = 0.6$

$\therefore P(150 < X < 170) = P(-0.4 < z < 0.6)$

$= A(0.6) + A(-0.4)$

$= 0.2258 + 0.1554$

$= 0.3812$



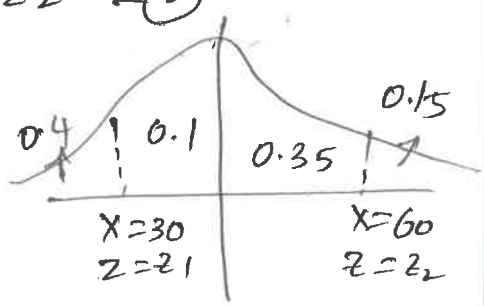
(3) The marks obtained in statistics in a certain exam found to be normally distributed. If 15% of the students ≥ 60 marks, 40% < 30 marks, find mean & S.D

Sol:

$$P(x < 30) = 0.4 \quad , \quad P(x \geq 60) = 0.15$$

when $x = 30$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{30-\mu}{\sigma} = -z_1$ — (1)

when $x = 60$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{60-\mu}{\sigma} = z_2$ — (2)



$\therefore \quad P(0 < z < z_1) = P(-z_1 < z < 0)$

$= 0.5 - 0.4$

$= 0.1$

& $P(0 < z < z_2) = 0.5 - 0.15 = 0.35$

from normal tables $z_1 = 0.25$ & $z_2 = 1.04$

$\therefore$ (1) $\Rightarrow$ $\dfrac{30-\mu}{\sigma} = -0.25$ & $\dfrac{60-\mu}{\sigma} = 1.04$

$30 - \mu = -0.25\sigma \qquad 60 - \mu = 1.04\sigma$

Solve above eqn's

$\therefore \quad \mu = 35.81, \quad \sigma = 23.26$

(5) The mean Voltage of a battery is '15' and Standard deviation 0.2. Find the probability that four such batteries connected in Series will have a Combined voltage of 60.8 (or) more Volts

Sol: Given that

$\mu = 15 \qquad \sigma = 0.2$

~~To find $p(X \geq 60.8)$~~

~~When $x = 60.8$ $\quad z = \frac{60.8 - 15}{0.2}$~~

Let mean voltage of a batteries 1,2,3,4 be $\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4$

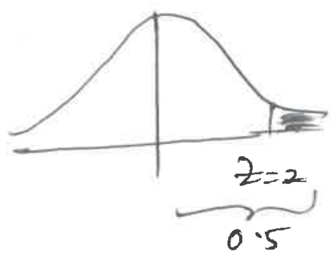
The mean of the series of the four batteries connected is

$\mu(\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \bar{x}_4) = \mu(\bar{x}_1) + \mu(\bar{x}_2) + \mu(\bar{x}_3) + \mu(\bar{x}_4) = 15+15+15+15 = 60$

$\sigma(\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \bar{x}_4) = \sqrt{\sigma^2(\bar{x}_1) + \sigma^2(\bar{x}_2) + \sigma^2(\bar{x}_3) + \sigma^2(\bar{x}_4)} = \sqrt{4(0.2)^2} = 0.4$

Let 'x' be the Combined Voltage of the Series

When $x = 60.8 \quad z = \frac{\bar{x} - \mu}{\sigma} = \frac{60.8 - 60}{0.4} = 2$



so $p(x \geq 60.8) = p(z \geq 2)$

$= 0.5 - A(2)$

$= 0.5 - 0.4772$

$= 0.0228.$

Mean of Normal distribution

Consider the Normal distribution with μ, σ as the parameters

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

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

$$= \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

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

$$x = \mu + \sigma z$$

$$dx = \cancel{0} + \sigma dz$$

$$= \frac{1}{\cancel{\sigma}\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu + \sigma z)\, e^{-\frac{1}{2}(z)^2}\, \cancel{\sigma} dz$$

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

$e^{-\frac{z^2}{2}}$ is even function $\qquad z e^{-\frac{z^2}{2}}$ odd function

$$\left[\int_{-\infty}^{\infty} f(x)\, dx = 2\int_{0}^{\infty} f(x)\, dx \quad f(x) \text{ is even} \atop \qquad = 0 \qquad\qquad f(x) \text{ is odd} \right]$$

$$= \frac{\mu}{\sqrt{2\pi}} \, 2 \int_0^\infty e^{-\frac{z^2}{2}} dz + 0$$

$$\frac{z^2}{2} = t$$

$$z^2 = 2t$$

$$z = \sqrt{2}\sqrt{t}$$

$$dz = \sqrt{2} \, \frac{1}{2} t^{\frac{1}{2}-1} dt$$

$$dz = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{t}} dt$$

$$= \frac{\mu}{\sqrt{2\pi}} \, 2 \int_0^\infty e^{-t} \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{t}} dt$$

$$= \frac{2\mu}{\cancel{\sqrt{2}}\sqrt{\pi}\cancel{\sqrt{2}}} \int_0^\infty e^{-t} t^{-\frac{1}{2}} dt$$

$$= \frac{\mu}{\sqrt{\pi}} \left(\Gamma\frac{1}{2} \right)$$

$$n-1 = -\frac{1}{2}$$

$$n = 1 - \frac{1}{2}$$

$$= \frac{1}{2}$$

$$\left[\int_0^\infty e^{-x} x^{n-1} dx = \Gamma\frac{1}{2} \right)$$

$$= \frac{\mu}{\cancel{\sqrt{\pi}}} \cancel{\sqrt{\pi}}$$

$$= \mu.$$

Variance of Normal distribution

$$Var(X) = E(X^2) - [E(X)]^2 \qquad (E(X) = \mu)$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$$

$$= \int_{-\infty}^{\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

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

$$x = \mu + \sigma z$$

$$dx = \sigma dz$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu + \sigma z)^2 e^{-\frac{1}{2}(z^2)} \sigma dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu^2 + \sigma^2 z^2 + 2\mu\sigma z) e^{-\frac{z^2}{2}} dz$$

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

$e^{-\frac{z^2}{2}}$ is even function

$z^2 e^{-\frac{z^2}{2}}$ even function

$z e^{-z^2/2}$ odd function.

$$= \frac{\mu^2}{\sqrt{2\pi}} 2\int_0^\infty e^{-\frac{z^2}{2}} dz + \frac{\sigma^2}{\sqrt{2\pi}} 2\int_0^\infty z^2 e^{-\frac{z^2}{2}} dz + 0$$

Let $\frac{z^2}{2} = t$

$$z^2 = 2t$$

$$z = \sqrt{2}\sqrt{t}$$

$$dz = \sqrt{2}\,\frac{1}{2} t^{\frac{1}{2}-1} dt$$

$$dz = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{t}} dt$$

$$= \frac{\mu^2}{\sqrt{2\pi}} 2\int_0^\infty e^{-t} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{t}} dt + \frac{\sigma^2}{\sqrt{2\pi}} 2\int_0^\infty 2t\, e^{-t} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{t}} dt$$

$$= \frac{\cancel{2}\mu^2}{\cancel{\sqrt{2}}\sqrt{2}\sqrt{\pi}} \int_0^\infty e^{-t} t^{-\frac{1}{2}} dt + \frac{\cancel{4}\sigma^2}{\cancel{\sqrt{2}}\sqrt{2}\sqrt{\pi}} \int_0^\infty e^{-t} t^{\frac{1}{2}} dt$$

$n-1 = -\frac{1}{2}$ $\quad$ $n-1 = \frac{1}{2}$

$n = \frac{1}{2}$ $\quad$ $n = \frac{3}{2}$

$$= \frac{\mu^2}{\sqrt{\pi}} \left(\Gamma\tfrac{1}{2}\right) + \frac{2\sigma^2}{\sqrt{\pi}} \left(\Gamma\tfrac{3}{2}\right)$$

$$= \frac{\mu^2}{\cancel{\sqrt{\pi}}} (\cancel{\sqrt{\pi}}) + \frac{\cancel{2}\sigma^2}{\cancel{\sqrt{\pi}}} \frac{1}{\cancel{2}} \cancel{\sqrt{\pi}}$$

$E(x^2) = \mu^2 + \sigma^2$

$\mathrm{var}(x) = (\mu^2 + \sigma^2) - \mu^2$

$\mathrm{Var}(x) = \sigma^2$

→ A population consists of five numbers 2, 3, 6, 8 and 11 Consider all possible samples of size two which can be drawn with replacement from this population. Find

i) the mean of the population.

(ii) The standard deviation of the population

(iii) The mean of the sampling distribution of means

(iv) The standard deviation of the sampling distributions of means (I.e the standard error of means)

Sol:-

a) Mean of the population is given by

$$\mu = \frac{2+3+6+8+11}{5} = \frac{30}{5} = 6$$

b) Variance of the population (σ^2) is given by

$$\sigma^2 = \Sigma \frac{(x_i - \bar{x})^2}{n}$$

$$= \frac{(2-6)^2 + (3-6)^2 + (6-6)^2 + (8-6)^2 + (11-6)^2}{5}$$

$$= \frac{16+9+0+4+25}{5} = 10.8$$

$$\sigma = \sqrt{10.8} =$$

$$= 3.29$$

c) Sampling with replacement (infinite population)

The total no. of samples with replacement is

$N^n = 5^2 = 25$ sample of size $= 2$

Here $N =$ population size

$n =$ sample size

The 25 samples

$$\begin{Bmatrix} (2,2) & (2,3) & (2,6) & (2,8) & (2,11) \\ (3,2) & (3,3) & (3,6) & (3,8) & (3,11) \\ (6,2) & (6,3) & (6,6) & (6,8) & (6,11) \\ (8,2) & (8,3) & (8,6) & (8,8) & (8,11) \\ (11,2) & (11,3) & (11,6) & (11,8) & (11,11) \end{Bmatrix}$$

Now mean of each of these 25 samples

The sample means are

$$\begin{Bmatrix} 2 & 2.5 & 4 & 5 & 6.5 \\ 2.5 & 3 & 4.5 & 5.5 & 7 \\ 4 & 4.5 & 6 & 7 & 8.5 \\ 5 & 5.5 & 7 & 8 & 9.5 \\ 6.5 & 7 & 8.5 & 9.5 & 11 \end{Bmatrix}$$

$$\mu_{\bar{x}} = \frac{\text{Sum of all sample means}}{25} = \frac{150}{25} = 6$$

So $\mu_{\bar{x}} = \mu$

d) $\sigma_{\bar{x}}^2 = \frac{(2-6)^2 + \cdots + (11-6)^2}{25} = \frac{135}{25} = 5.40$

and $\sigma_{\bar{x}} = \sqrt{5.40} = 2.32$

Clearly for finite population involving sampling with replacement (Infinite population)

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{(3.29)^2}{2} = 2.32$$

Solution of above problem with out Replacement (finite population)

Sol

i) $\mu = \frac{2+3+6+8+11}{5} = \frac{30}{5} = 6$

ii) $\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{n}$

$= \frac{(2-6)^2 + (3-6)^2 + (6-6)^2 + (8-6)^2 + (11-6)^2}{5}$

$\sigma^2 = 10.8$ so $\sigma = 3.29$

iii) Sampling with out replacement

The total no of samples with out replacement is ${}^{N}C_{n} = {}^{5}C_{2} = 10$

The 10 samples are

$$\left\{ \begin{array}{llll} (2,3) & (2,6) & (2,8) & (2,11) \\ (3,6) & (3,8) & (3,11) & \\ (6,8) & (6,11) & & \\ (8,11) & & & \end{array} \right\}$$

The Corresponding Sample means are

$$\begin{Bmatrix} 2.5 & 4 & 5 & 6.5 \\ 4.5 & 5.5 & 7 & \\ 7 & 8.5 & & \\ 9.5 & & & \end{Bmatrix}$$

The mean of the Sampling distribution of means

$$\mu_{\bar{x}} = \frac{2.5+4+5+6.5+4.5+5.5+7+7+8.5+9.5}{10} = 6$$

So $\mu_{\bar{x}} = \mu$

iv) The variance of Sampling distributions of means

$$\sigma_{\bar{x}}^2 = \frac{(2.5-6)^2+(4-6)^2+ \cdots +(9.5-6)^2}{10} = 4.05$$

$$\sigma_{\bar{x}} = 2.01$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{10.8}{2}\left(\frac{5-2}{5-1}\right) = 4.05$$

for Samping with out Replacement

UNIT- III

Continous probability distribution :-

Normal distribution :-

Normal distribution was discovered by Karl Friedrich gauss so that it is also called Gaussian distribution.

The Normal distribution is limiting case of Binomial distribution under the conditions.

1) No. of trials n is very large. i.e, $n \to \infty$

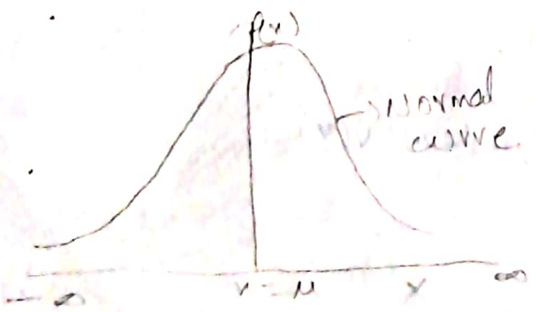
ii) Neither p nor q is very small.

A continous R.V 'x' is said to have normal distribution if its density function is given by

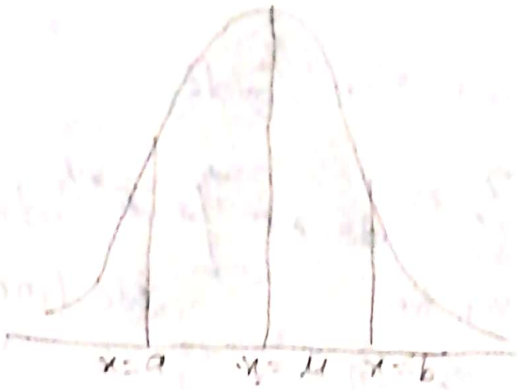
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

where μ, σ are mean & standard deviations of X

The graph of the R.V 'x' and density function $f(x)$ is called Normal curve. The total area bounded by Normal curve and x axis is equal to 1.



$$\int_{-\infty}^{\infty} f(x)\,dx = 1 = 100\%.$$



$P(a \le x \le b) = \int_a^b f(x)\,dx$ = Area under the curve b/w lines $x=a$, $x=b$.

Normal Distribution case of B.D:-

Formulae:-

1) $\int_{-a}^{a} f(x)\,dx = \begin{cases} 2\int_0^a f(x)\,dx & , f(x) \text{ even} \\ 0 & , f(x) \text{ odd} \end{cases}$

2) if $\int_a^b f(x)\,dx = 0$, then $a=b$ $\{f(x) \neq 0\}$

3) $\int_0^{\infty} e^{-x^2/2}\,dx = \sqrt{\pi/2}$

4) $\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \sqrt{2\pi}$.

5) i) $\Gamma n = \int_0^{\infty} e^{-x} \cdot x^{n-1}\,dx$.

ii) $\Gamma n = (n-1)\Gamma(n-1)$ if n is +ve function.

iii) $\Gamma 1/2 = \sqrt{\pi}$ $\quad \Gamma 1 = 1$.

Mean of Normal distribution:-

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

$E(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,dx$.

let $\frac{x-\mu}{\sigma} = z \Rightarrow \boxed{x = \mu + \sigma z}$

$\frac{dx}{\sigma} = dz$ $\boxed{\therefore dx = \sigma\,dz}$

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

$$= \frac{1}{\sqrt{2\pi}} \left\{ \mu \int_{-\infty}^{\infty} e^{-z^2/2} dz + \sigma \int_{-\infty}^{\infty} z \cdot e^{-z^2/2} \cdot dz \right\}.$$

even func $\qquad\qquad$ odd func.

$$E(x) = \frac{1}{\sqrt{2\pi}} \left\{ 2\mu \int_{0}^{\infty} e^{-z^2/2} dz + 0 \right\}.$$

$$E(x) = \frac{2\mu}{\sqrt{2\pi}} \cdot \sqrt{\frac{\pi}{2}}$$

$$= \frac{2\mu\sqrt{\pi}}{\sqrt{2} \cdot \sqrt{2}\sqrt{\pi}}$$

$$\boxed{\therefore E(x) = \mu.}$$

Variance of N.D :-

$$V(x) = E[(x-\mu)^2] = E(x^2) - \mu^2$$

$$E[(x-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx.$$

$$V(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2} dx.$$

let $z = \frac{x-\mu}{\sigma} \Rightarrow x - \mu = \sigma z$

$dz = \frac{dx}{\sigma}$.

$dx = \sigma dz$.

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

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

$$= \frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty} z^2 \cdot e^{-z^2/2} dz.$$

let $z^2/2 = t$

$z^2 = 2t$.

$$z = \sqrt{2}\sqrt{t}.$$

$$dz = \sqrt{2} \cdot \frac{1}{2\sqrt{t}} dt$$

$$dz = \frac{1}{\sqrt{2}\sqrt{t}} dt.$$

$$V(X) = \frac{2\sigma^2}{\sqrt{2\pi}} \int_0^\infty e^{-t} 2t \cdot \frac{1}{\sqrt{2}\sqrt{t}} dt$$

$$= \frac{4\sigma^2}{2\sqrt{\pi}} \int_0^\infty e^{-t} \cdot t^{1-1/2} dt.$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \int_0^\infty e^{-t} t^{1/2} dt.$$

$$\Rightarrow \frac{2\sigma^2}{\sqrt{\pi}} \int_0^\infty e^{-t} \cdot t^{3/2-1} dt.$$

$$\Rightarrow \frac{2\sigma^2}{\sqrt{\pi}} \Gamma 3/2$$

$$\Rightarrow \frac{2\sigma^2}{\sqrt{\pi}} \cdot \frac{1}{2} \Gamma 1/2$$

$$\Rightarrow \frac{2\sigma^2}{\sqrt{\pi}} \cdot \frac{1}{2} \sqrt{\pi}$$

$$\therefore V(X) = \sigma^2.$$

Mode of N.D :-

Mode is a value of 'x' for which $f(x)$ is maximum.

We get mode by $f'(x)=0$, $f''(x)<0$ at mode.

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

d.w.r.t to 'x'

$$f'(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2} \cdot \frac{d}{dx}\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

$$f'(x) = -\frac{1}{2}[f(x)] \cdot 2\left(\frac{x-\mu}{\sigma}\right)\frac{1}{\sigma}$$

$$f'(x) = -f(x)\left(\frac{x-\mu}{\sigma^2}\right) \rightarrow (1)$$

$$\therefore f'(x) = 0.$$

$f(x)\left(\frac{x-\mu}{\sigma^2}\right)=0$

$\frac{x-\mu}{\sigma^2}=0$

$\boxed{\therefore x=\mu}$

d.w.r.t to 'x'

$f''(x)=-\left[f'(x)\left(\frac{x-\mu}{\sigma^2}\right)+f(x)\frac{1}{\sigma^2}\right]$

at $x=\mu$,

$f''(\mu)=-\left[0+f(\mu)\frac{1}{\sigma^2}\right]$

$f''(\mu)=-\frac{1}{\sigma\sqrt{2\pi}}\cdot\frac{1}{\sigma^2}<0$

$\therefore$ Mode $=\mu$.

Median of N.D :-

Let 'M' be median of N.D' $x\in[-\infty,\infty]$.

$\int_{-\infty}^{M} f(x)dx=\int_{M}^{\infty} f(x)dx=1/2$.

$\int_{-\infty}^{m} f(x)dx=1/2$.

Suppose, $-\infty<\mu<m$,

$\int_{-\infty}^{\mu} f(x)dx+\int_{\mu}^{m} f(x)dx=1/2 \rightarrow (1)$

Consider,

$\int_{-\infty}^{\mu} f(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\mu} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}dx$.

let $\frac{x-\mu}{\sigma}=z$ if $x=-\infty \Rightarrow z=-\infty$

$\frac{dx}{\sigma}=dz$. if $x=\mu \Rightarrow z=0$.

$dx=\sigma dz$

$$\Rightarrow \int_{-\infty}^{\mu} f(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \cdot \int_{-\infty}^{0} e^{-z^2/2} (\sigma\, dz)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-z^2/2} dz$$

$$\Rightarrow \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-z^2/2} dz$$

$$\Rightarrow \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\pi}{2}} \Rightarrow 1/2 .$$

$\int_0^{\infty} e^{-x^2/2} dx = \sqrt{\frac{\pi}{2}}$

if $\int_a^b f(x)\,dx = 0$ $(a=b)$

9, fr from ①,

$$\frac{1}{2} + \int_{\mu}^{m} f(x)\,dx = 1/2 .$$

$$\int_{\mu}^{m} f(x)\,dx = 1/2 - 1/2 = 0$$

then $m = \mu$.

$\boxed{\therefore \text{median} = \mu}$.

Mean deviation :-

$$M.D = \int_{-\infty}^{\infty} |x-\mu|\, f(x)\,dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

let $\frac{x-\mu}{\sigma} = z \Rightarrow x-\mu = \sigma z$

$\frac{dx}{\sigma} = dz \Rightarrow dx = \sigma dz$.

$\Gamma n = \int_0^{\infty} e^{-x} x^{n-1} dx$

$$M.D = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} |\sigma z|\, e^{-z^2/2} (\sigma dz)$$

$\Gamma\frac{1}{2} = \sqrt{\pi}$, $\Gamma 1 = 1$

$$= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \underbrace{|z| \cdot e^{-z^2/2}}_{\text{even fn.}} dz .$$

$$= \frac{2\sigma}{\sqrt{2\pi}} \int_{0}^{\infty} |z| \cdot e^{-z^2/2} dz .$$

$$\Rightarrow \frac{\sqrt{2}\sigma}{\sqrt{\pi}} \int_0^\infty z \cdot e^{-z^2/2} \, dz$$

let $z^2/2 = t \Rightarrow z^2 = 2t$

$z = \sqrt{2} \cdot \sqrt{t}$

$dz = \sqrt{2} \cdot \frac{1}{2\sqrt{t}} dt$

$dz = \frac{1}{\sqrt{2} \cdot \sqrt{t}} dt$

$$\Rightarrow \frac{\sqrt{2}\sigma}{\sqrt{\pi}} \int_0^\infty \sqrt{2} \cdot \sqrt{t} \cdot e^{-t} \cdot \frac{1}{\sqrt{2}\sqrt{t}} \, dt.$$

$$\Rightarrow \frac{\sqrt{2}\sigma}{\sqrt{\pi}} \int_0^\infty e^{-t} \cdot t^0 \, dt.$$

$\Gamma\frac{1}{2} = \sqrt{\pi}$

$\Gamma 1 = 1$

$$\Rightarrow \frac{\sqrt{2}\sigma}{\sqrt{\pi}} \int_0^\infty e^{-t} \cdot t^{1-1} \, dt.$$

$$\Rightarrow \frac{\sqrt{2}\sigma}{\sqrt{\pi}} \cdot \Gamma(1)$$

$$\Rightarrow \sqrt{\frac{2}{\pi}}\,\sigma \Rightarrow \frac{4}{5}\sigma.$$

Chief characteristics of N.D :-

1) The graph of Normal distribution $y = f(x)$ in XY plane is called normal curve.

ii) The curve is Bell shapped and symmetrical about mean line $x = \mu$. and the 2 tails on the right and left sides of mean extends to infinity.

3) i) The area under the Normal curve represents total population.

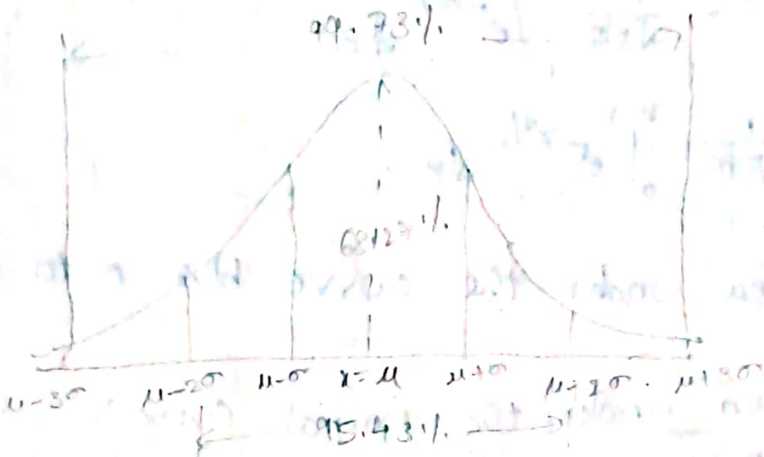
ii) Mean, median & mode of distribution are coincide at $x = \mu$.

i.e; mean = median = mode.

Scanned by CamScanner

4) X-axis is asymptote to the curve. And normal curve is unimodel.

5)



Area b/w $\mu-\sigma$ & $\mu+\sigma$ is 68.27%

i.e; $P(\mu-\sigma < x < \mu+\sigma) = 68.27\%$

$P(\mu-2\sigma < x < \mu+2\sigma) = 95.43\%$

$P(\mu-3\sigma < x < \mu+3\sigma) = 99.73\%$.

6) Normal variable:-

The variable $\boxed{z = \frac{x-\mu}{\sigma}}$ is called normal variable.

Standard Normal variables:-

Normal variable z is called standard normal variable if $\mu=0, \sigma=1$.

Area under the Normal Curve:-

If the R.V 'x' lies b/w μ & x_1, then probability of 'x' lies b/w μ & x_1 is

$$P(\mu < x < x_1) = \int_{\mu}^{x_1} f(x)\,dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu}^{x_1} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2} dx.$$

let $z = \frac{x-\mu}{\sigma}$ | if $x=\mu \Rightarrow z=0$

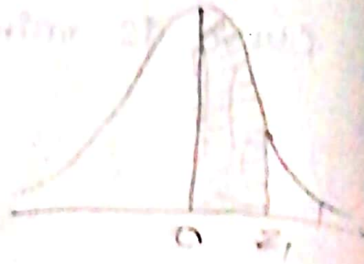
$dz = \frac{dx}{\sigma}$ | if $x=x_1 \Rightarrow z_1 = \frac{x_1-\mu}{\sigma}$.

Scanned by CamScanner

$$P(0<Z<z_1) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^{z_1} e^{-z^2/2} (\sigma\, dz).$$

$$P(0<z<z_1) = \frac{1}{\sqrt{2\pi}} \int_0^{z_1} e^{-z^2/2}\, dz$$

$$A(z_1) = \frac{1}{\sqrt{2\pi}} \int_0^{z_1} e^{-z^2/2}\, dz.$$



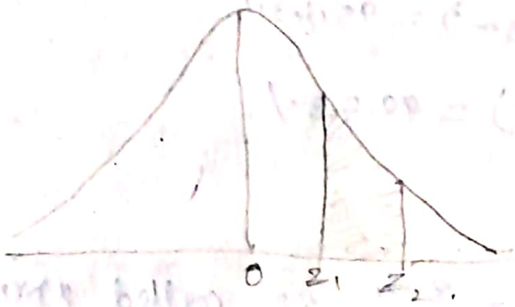
$A(z_1)$ is Area under the curve b/w 0 to z_1

Finding the area under the Normal curve :-

1) for $P(z_1 < z < z_2)$:-

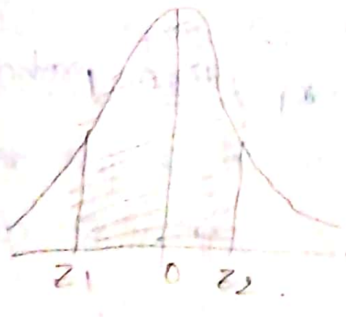
case - 1 :- if $z_1 > 0$, $z_2 > 0$.

for $P(z_1 < z < z_2)$:-



$$P(z_1 < z < z_2) = P(0 < z < z_2) - P(0 < z < z_1)$$
$$= A(z_2) - A(z_1);$$

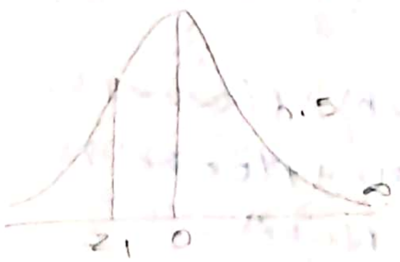
case - 2 :- if $z_1 < 0$, $z_2 > 0$



$$P(z_1 < z < z_2) = P(z_1 < z < 0) + P(0 < z < z_2)$$
$$= P(0 < z < z_1) + P(0 < z < z_2)$$
$$= A(z_1) + A(z_2)$$

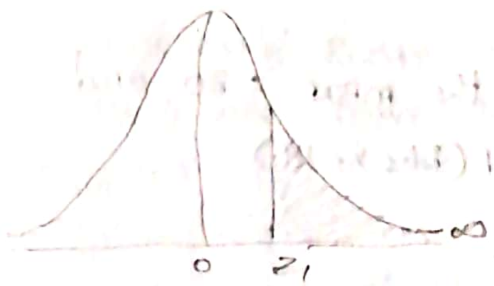
2) for $P(z > z_1)$:-

Case - I :- if $z_1 < 0$.



$$P(z > z_1) = 0.5 + P(z_1 < z < 0)$$
$$= 0.5 + P(0 < z < z_1)$$
$$= 0.5 + A(z_1).$$

Case - II :- if $z_1 > 0$



$$P(z > z_1) = 0.5 - P(0 < z < z_1)$$
$$= 0.5 - A(z_1).$$

1) Normally distributed variable with mean $(\mu) = 1$, standard deviation $(\sigma) = 3$. Find (i) $P(3.43 \le x \le 6.19)$

(ii) $P(-1.43 \le x \le 6.19)$.

i) $P(3.43 \le x \le 6.19)$

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

if $x = 3.43 \Rightarrow z = \frac{3.43 - 1}{3} = 0.81$

if $x = 6.19 \Rightarrow z = \frac{6.19 - 1}{3} = \frac{5.19}{3} = 1.73.$

$$P(0.81 \le z \le 1.73) = P(0 < z < 1.73) - P(0 < z < 0.81)$$
$$= A(1.73) - A(0.81)$$
$$= 0.4582 - 0.2910$$
$$= 0.1672$$

ii) $P(-1.43 \le z \le 6.19)$.

If $x = -1.43 \Rightarrow z = \frac{-1.43-1}{3} \Rightarrow \frac{-2.43}{3} \Rightarrow -0.81$.

If $x = 6.19 \Rightarrow z = \frac{6.19-1}{3} \Rightarrow \frac{5.19}{3} \Rightarrow 1.73$.

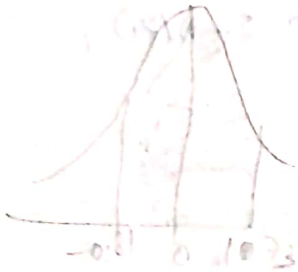
$P(-0.81 < z < 1.73) = P(-0.81 < z < 0) + P(0 < z < 1.73)$

$= P(0 < z < 0.81) + P(0 < z < 1.73)$

$= A(0.81) + A(1.73)$

$= 0.2910 + 0.4582$

$= 0.7492$



2) If x is normal variate if the mean $= 30$ and Std. deviation 5 then find i) $P(26 \le x \le 40)$
ii) $P(x \ge 45)$ iii) $P(x \le 45)$

i) Mean $(\mu) = 30$

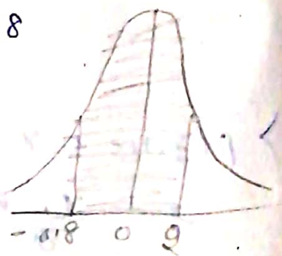
S.D $(\sigma) = 5$.

$P(26 \le x \le 40)$

If $x = 26 \Rightarrow z = \frac{26-30}{5} \Rightarrow -4/5 \Rightarrow -0.8$

If $x = 40 \Rightarrow z = \frac{40-30}{5} \Rightarrow 10/5 \Rightarrow 2$.

$P(26 \le x \le 40) \Rightarrow P(-0.8 \le z \le 2)$



$\Rightarrow P(-0.8 \le z \le 0) + P(0 \le z \le 2)$

$\Rightarrow P(0 \le z \le 0.8) + P(0 \le z \le 2)$

$\Rightarrow A(0.8) + A(2)$

$\Rightarrow 0.2881 + 0.4772$

$\Rightarrow 0.7653$.

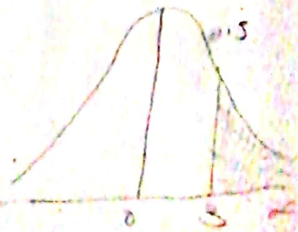
ii) $P(x \ge 45)$

If $x = 45 \Rightarrow z = \frac{45-30}{5} \Rightarrow \frac{15}{3} \Rightarrow 3$.

$P(x \ge 45) = P(z \ge 3)$

$\Rightarrow 0.5 - P(0 \le z \le 3) \Rightarrow 0.5 - A(3)$

$\Rightarrow 0.5 - 0.4987 \Rightarrow 0.0013$



Scanned by CamScanner

iii) $P(x \le 45)$

if $x = 45 \Rightarrow z = \frac{45-30}{5} \Rightarrow 3$.

$P(x \le 45) \Rightarrow 0.5 + P(0 \le z \le 3)$

$\Rightarrow 0.5 + A(3)$

$\Rightarrow 0.5 + 0.4987$.

$\Rightarrow 0.9987$.

3) In a sample of 1000 cases, mean ~~is obtai~~ of certain test is 14, and std deviation 2.5. Assume the distribution to be normal. Find i) how many students scored b/w 12 & 15.

ii) How many score above 18.

iii) How many score below 18.

Mean(μ) = 14

S.D (σ) = 2.5

Let x = score of student.

i) $P(12 < x < 15)$

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

if $x = 12 \Rightarrow z = \frac{12-14}{2.5} \Rightarrow \frac{-2}{2.5} \Rightarrow -0.8$.

if $x = 15 \Rightarrow z = \frac{15-14}{2.5} \Rightarrow \frac{1}{2.5} \Rightarrow \frac{10}{25} \Rightarrow 0.4$.

$P(-0.8 < z < 0.4) = P(-0.8 < z < 0) + P(0 < z < 0.4)$

$= A(-0.8) + A(0.4)$

$= 0.2881 + 0.1554$

$= 0.4435$.

No. of students score b/w 12 & 15 is

$\Rightarrow 0.4435 \times 1000$

$\Rightarrow 443.5$

$\Rightarrow 444$ students.

ii) $P(x > 18)$

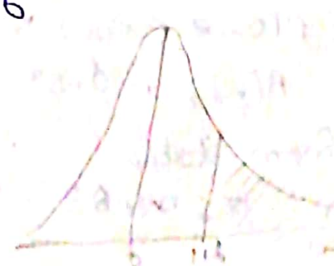
if $x = 18 \Rightarrow z = \frac{18-14}{2.5} \Rightarrow \frac{4}{2.5} \Rightarrow 1.6$.

$P(z > 1.6) = 0.5 - P(0 < z < 1.6)$

$= 0.5 - A(1.6)$

$= 0.5 - 0.4452$

$\Rightarrow 0.0548$



Scanned by CamScanner

No. of students score above 18 $\Rightarrow 0.0548 \times 1000$

$\Rightarrow 54.8$

$\Rightarrow 55.$

iii) $P(X<18)$.

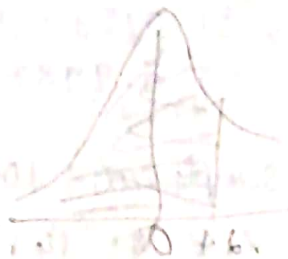
if $X=18 \Rightarrow z=1.6$.

$P(z<1.6) = 0.5 + P(0<z<1.6)$

$= 0.5 + A(1.6)$

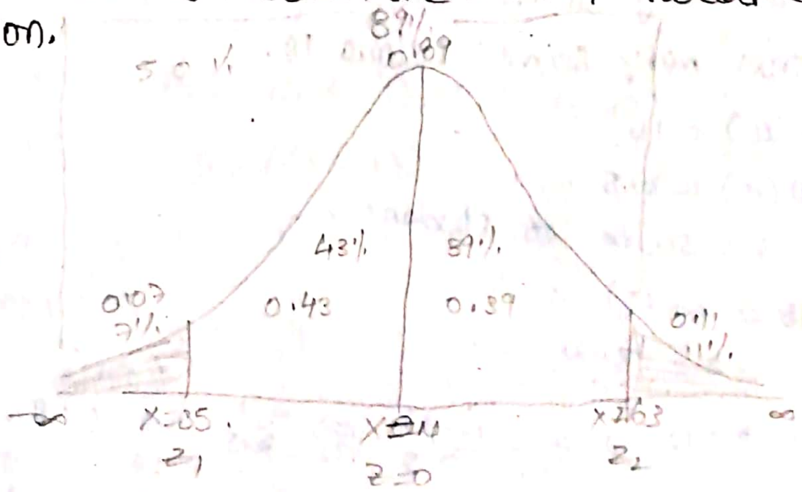
$\Rightarrow 0.5 + 0.4452$

$\Rightarrow 0.9452.$



***
***

4) In a Normal dist. 7% of items under 35 and 89% of items under 63. Determine mean & variance of distribution.



$P(X<35) = 7\% = 0.07$

$P(X<63) = 89\% = 0.89$

$P(X\geq 63) = 1 - P(X<63)$

$= 1 - 0.89 \Rightarrow 0.11$

$z = \dfrac{x-\mu}{\sigma}$.

if $x=35 \Rightarrow \dfrac{35-\mu}{\sigma} \Rightarrow -z_1$ (say)

if $x=63 \Rightarrow \dfrac{63-\mu}{\sigma} \Rightarrow z_2$ (say) $\Big\} \rightarrow ①$

from diagram,

$P(0<z<z_2) = 0.39$

$A(z_2) = 0.39$

from table

$z_2 = 1.23$

Scanned by CamScanner

$P(-z_1 < z < 0) = 0.43.$

$A(-z_1) = 0.43.$

$A(z_1) = 0.43.$

from table

$z_1 = 1.48.$

from ①.

$$\frac{35}{\sigma} - \frac{\mu}{\sigma} = -1.48.$$

$$-\frac{63}{\sigma} - \frac{\mu}{\sigma} = 1.23$$

$$\frac{-28}{\sigma} = -2.71.$$

$$\boxed{\therefore \sigma = 10.33.}$$

from ①

$$\frac{35-\mu}{\sigma} = -1.48.$$

$$\frac{35-\mu}{10.33} = -1.48$$

$$35-\mu = -1.48 \times 10.33.$$

$$\boxed{\mu = 50.28.}$$

5) In Normal distr of 31% items under 45 & 8% of items over 64. Find Mean & variance of distribution.

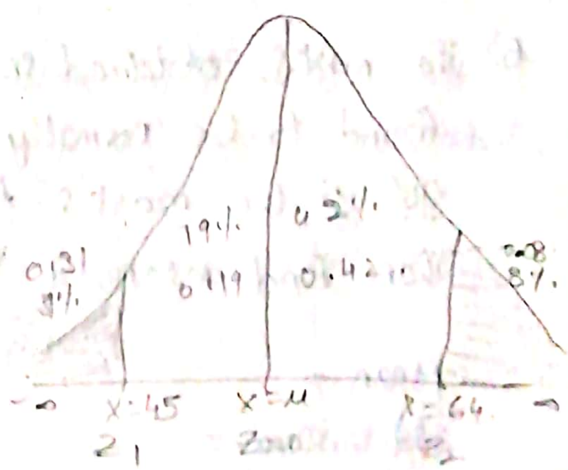
$P(x < 45) = 0.31$

$P(x > 64) = 8\% = 0.08$

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

if $x = 45 \Rightarrow \frac{45-\mu}{\sigma} = -z_1$ (say).

if $x = 64 \Rightarrow \frac{64-\mu}{\sigma} \Rightarrow z_2$ (say) $\Big\} \to ①$



from diagram

$P(0 < z < z_2) = 0.42$

$A(z_2) = 0.42$

from table

$$\boxed{z_2 = 1.41.}$$

$P(-z_1 < Z < 0) = 0.19$

$A(-z_1) = 0.19$

$A(z_1) = 0.19$

From table

$z_1 = 0.5.$

From ①.

$$\frac{45}{\sigma} - \frac{\mu}{\sigma} = -0.50$$

$$\frac{64}{\sigma} - \frac{\mu}{\sigma} = 1.41$$

$$\frac{-19}{\sigma} = -1.91$$

$$\sigma = \frac{-19}{-1.91}$$

$$\boxed{\sigma = 9.94.}$$

From ①

$$\frac{45-\mu}{\sigma} = -0.5.$$

$$\frac{45-\mu}{9.94} = -0.5.$$

$45-\mu = -0.5 \times 9.94.$

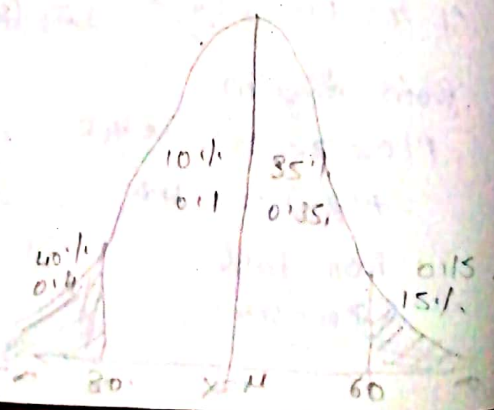
$\mu = 45 + 0.5 \times 9.94.$

$$\boxed{\mu = 55.4}$$

6) The marks obtained in Statistics in a certain examination found to be normally distributed. If 15% of student got ≥ 60 marks, 40% of students got < 30 marks. Then find Mean & Variance of distribution.

Mean =

Std variance =



7) The marks obtained in mathematics by 1000 students is normally distributed with mean 78% and std-deviation 11%. determine (i) how many students got marks above 90%. (ii) what was the highest mark obtained by lowest 10% of students (iii) within what limits did middle of 90% of students lie.

8) Given that the mean height of students in a college is 155 cm and S.D 15. what is the probability that mean height of 36 students is less than 157 cm.

9) If X is normally distributed with mean 2 & variance S.D is 0.1, then find $P(|x-2| \geq 0.01)$

7) Let x = Marks percentage of students.

Mean $(\mu) = 78\% = 0.78$.

S.D $(\sigma) = 11\% = 0.11$.

i) $P(x > 90\%)$

if $x = 90\% = 0.9 \Rightarrow z = \frac{x-\mu}{\sigma} \Rightarrow \frac{0.9-0.78}{0.11} \Rightarrow 1.09$.

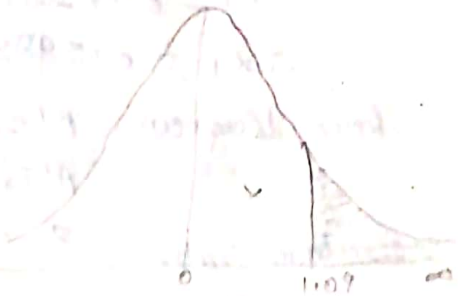
$P(x > 0.9) = P(z > 1.09)$

$\Rightarrow 0.5 - P(0 < z < 1.09)$

$\Rightarrow 0.5 - A(1.09)$

$\Rightarrow 0.5 - 0.3621$

$\Rightarrow 0.138$.



No. of students got marks above 90% $\Rightarrow 0.138 \times 1000$

$\Rightarrow$ 138 students.

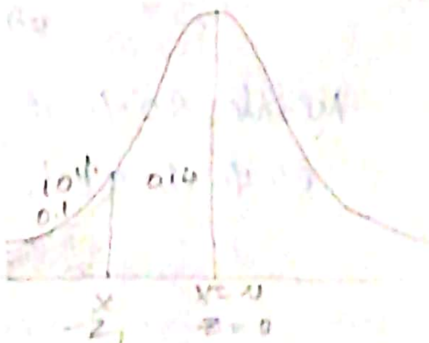
ii) $P(-z_1 < z < 0) = 0.4$.

$A(-z_1) = 0.4$

from tables,

$-z_1 = 1.29$

$\boxed{z_1 = -1.29}$



$\frac{x_1 - \mu}{\sigma} = -1.29$.

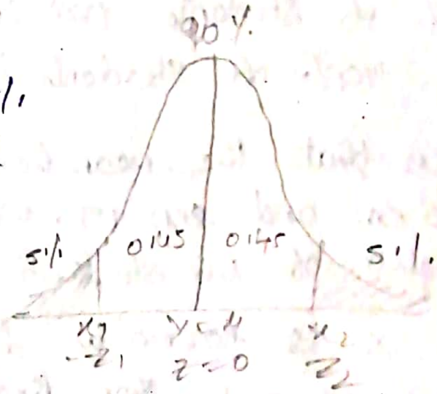
$\frac{x_1 - 0.78}{0.11} = -1.29$.

$x_1 - 0.78 = -1.29 \times 0.11$

$x_1 = 0.6381 \Rightarrow 63.81\% \cong 64\%$

Highest marks obtained by lowest 10% of student is 64%.

iii) Middle 90% means leaving 5% area on both sides of normal curve.



from diagram:-

$P(-z_1 < z < 0) = 0.45$

$A(-z_1) = 0.45$.

$-z_1 = 1.65$

$z_1 = -1.65$. (from Tables).

$\frac{x_1 - \mu}{\sigma} = -1.65$

$\frac{x_1 - 0.78}{0.11} = -1.65$

$x_1 - 0.78 = -1.65 \times 0.11$

$x_1 = 0.78 - (1.65 \times 0.11)$

$\therefore x_1 = 0.5985 \Rightarrow 59.85\% \Rightarrow 60\%$

from diagram, $P(0 < z < z_2) = 0.45$

$A(z_2) = 0.45$.

from Tables $z_2 = 1.65$

$\frac{x_2 - \mu}{\sigma} = 1.65$,

$\frac{x_2 - 0.78}{0.11} = 1.65$

$\therefore x_2 = 0.78 + (1.65 \times 0.11)$.

$\therefore x_2 = 0.9615 \Rightarrow 96.15\% \Rightarrow 96\%$

Middle 90% of student got marks b/w 60% & 90%.

8) Mean(μ) = 155 cm

S.D(σ) = 15 cm,

No. of students (n) = 36.

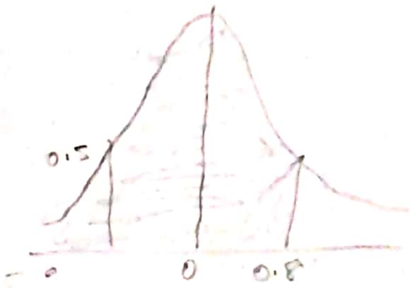
Sample mean ($\bar{x}$) = 157

$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$

To find $P(\bar{x} < 157)$

if $\bar{x} = 157 \Rightarrow z = \frac{157-155}{15/\sqrt{36}}$

$\Rightarrow \frac{2\times 6}{15} \Rightarrow \frac{12}{15} \Rightarrow 0.8$.



$P(z < 0.8) = 0.5 + P(0 < z < 0.8)$

$\Rightarrow 0.5 + A(0.8)$

$\Rightarrow 0.5 + 0.2881$

$\Rightarrow 0.7881$

9)

$\mu = 2$

$\sigma = 0.1$

To find $P(|x-2| \geq 0.01)$.

$|x| < a \Leftrightarrow -a < x < a$

$|x| > a \Leftrightarrow x \in (-\infty, -a] \cup [a, \infty)$

* $P(|x-2| \geq 0.01) = 1 - P(|x-2| < 0.01) \rightarrow (1)$.

$|x-2| < 0.01$

$-0.01 < (x-2) < 0.01$

$2 - 0.01 < x < 2 + 0.01$

$1.99 < x < 2.01$

if $x = 1.99 \Rightarrow z = \frac{1.99-2}{0.1} \Rightarrow \frac{-0.01}{0.1} \Rightarrow -0.1$

if $x = 2.01 \Rightarrow z = \frac{2.01-2}{0.1} \Rightarrow \frac{0.01}{0.1} \Rightarrow 0.1$

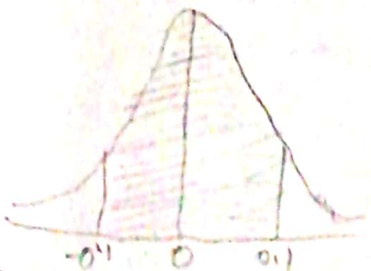
$P(|x-2| < 0.01) = P(1.99 < x < 2.01)$

$= P(-0.1 < z < 0.1)$

$= P(-0.1 < z < 0) + P(0 < z < 0.1)$

$= A(-0.1) + A(0.1)$

$= 2A(0.1) \Rightarrow 2(0.0398) \Rightarrow 0.0796$.



Normal approximation to Binomial distribution :-

Let X be a Random variable of Binomial distribution.
& n = no. of trails, p = prob. of success and
q = prob. of failure, $\mu = np$, $\sigma^2 = npq$
$\sigma = \sqrt{npq}$

To find prob. of x lies b/w x_1 & x_2
i.e., $P(x_1 \le x \le x_2)$

if $x = x_1 \Rightarrow z_1 = \dfrac{(x_1 - 1/2) - \mu}{\sigma}$

if $x = x_2 \Rightarrow z_2 = \dfrac{(x_2 + 1/2) - \mu}{\sigma}$

$P(x_1 < x < x_2) = P(z_1 < z < z_2)$.

Find the probability that out of 100 ~~students~~ patient b/w 84 & 95 inclusive will survive from heart operation given that the chances of survival is 0.9.

$P = 0.9 \quad n = 100$

$q = 1 - p \Rightarrow 1 - 0.9 = 0.1$

Let X = no. of survivals.

$\mu = np = (100)(0.9) = 90$

$\sigma = \sqrt{npq} \Rightarrow \sqrt{(90)(0.1)} \Rightarrow \sqrt{9} \Rightarrow 3$.

To find $P(84 \le x \le 95)$

if $x_1 = 84 \Rightarrow z_1 = \dfrac{(x_1 - 1/2) - \mu}{\sigma} \Rightarrow \dfrac{(84 - 1/2) - 90}{3} \Rightarrow -2.16$.

if $x_1 = 95 \Rightarrow z_2 = \dfrac{(x_2 + 1/2) - \mu}{\sigma} \Rightarrow \dfrac{(95 + 1/2) - 90}{3} = 1.83$.

$P(84 \le x \le 95) = P(-2.16 \le z \le 1.83)$

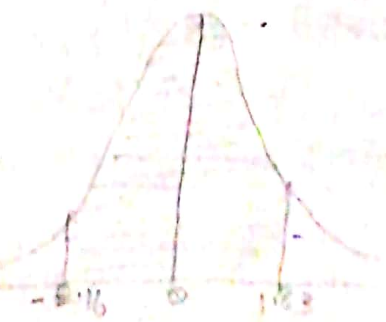
$\Rightarrow P(-2.16 \le z \le 0) + P(0 \le z \le 1.83)$.

$\Rightarrow A(-2.16) + A(1.83)$

$\Rightarrow A(2.16) + A(1.83)$

$\Rightarrow 0.4846 + 0.4064$

$\Rightarrow 0.952$.



2) 8 coins are tossed together find the probability that getting 1 to 4 heads in a single toss.

$p = 1/2 \qquad n = 8$

$q = 1-p \Rightarrow 1 - 1/2 = 1/2$

Let x = No. of heads.

$\mu = np = 8(1/2) = 4$

$\sigma = \sqrt{npq} \Rightarrow \sqrt{8(\frac{1}{2})(\frac{1}{2})} = \sqrt{2}$

To find $P(1 < x < 4)$.

3) Find the prob. of getting even number on the face in 3 to 5 times in throwing 10 dice together

$p = \frac{3}{6} \Rightarrow 1/2 \qquad n = 10$

$q = 1 - 1/2 \Rightarrow 1/2$

Let x = no of times getting even number

$\mu = np =$

$\sigma = \sqrt{npq}$

$P(\cancel{3} < x < 5)$.

Exponential Distribution :-

If X is a continuous R.V probability density function.

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x}, & \text{if } x \geq 0, \lambda > 0. \\ 0 & , \text{elsewhere.} \end{cases}$$

1) Mean :-

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

$\int uv\,dx = uv_1 - u'v_2 + u''v_3 - u'''v_4 \ldots$

$v_1 = \int v\,dx$, $v_2 = \int\int v\,dx\,dx$

$$E(x) = 0 + \int_0^{\infty} \underset{u}{x} \cdot \lambda \cdot \underset{v}{e^{-\lambda x}} dx$$

$$E(x) = \lambda \left\{ x\left(\frac{e^{-\lambda x}}{-\lambda}\right) - 1 \cdot \frac{e^{-\lambda x}}{\lambda^2} \right\}_0^{\infty}$$

$$E(x) = \lambda \{ [0-0] - [0 - 1/\lambda^2] \}.$$

$$E(x) = \text{mean} = 1/\lambda$$

2) Variance :-

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x) dx.$$

$$E(x^2) = 0 + \lambda \int_0^{\infty} \underset{u}{x^2} \underset{v}{e^{-\lambda x}} dx$$

$$\Rightarrow \lambda \left\{ x^2 \left(\frac{e^{-\lambda x}}{-\lambda}\right) - (2x)\left(\frac{e^{-\lambda x}}{\lambda^2}\right) + 2\left(\frac{e^{-\lambda x}}{-\lambda^3}\right) \right\}_0^{\infty}$$

$$\Rightarrow \lambda \{ [0+0+0] - [0+0-2/\lambda^3] \}.$$

$$E(x^2) = \lambda \cdot 2/\lambda^3 \Rightarrow 2/\lambda^2$$

$$V(x) = E(x^2) - \mu^2 \Rightarrow 2/\lambda^2 - 1/\lambda^2$$

$$\therefore V(x) = 1/\lambda^2$$

1) The length of time a person speaks over phone follows exponential distribution with mean 6. What is the probability that person will talk for (i) More than 8mins (ii) B/w 4 & 8 mins.

Mean $= 6 = 1/\lambda$

$\therefore \lambda = 1/6$.

$$f(x) = \begin{cases} 1/6 \, e^{-x/6}, & x \geq 0. \\ 0, & \text{otherwise}. \end{cases}$$

Let x = length of time.

i) $P(x > 8) = \int_8^\infty f(x)\,dx \Rightarrow \frac{1}{6}\int_8^\infty e^{-x/6}\,dx$

$\Rightarrow 1/6 \left[\frac{e^{-x/6}}{-1/6}\right]_8^\infty \Rightarrow -\{e^{-\infty} - e^{-8/6}\}. \Rightarrow e^{-4/3}$.

ii) $P(4 < x < 8) = \int_4^8 f(x)\,dx$

$\Rightarrow 1/6 \int_4^8 e^{-x/6}\,dx \Rightarrow 1/6 \left[\frac{e^{-x/6}}{-1/6}\right]_4^8$

$\Rightarrow e^{-4/6} - e^{-8/6}$

$\Rightarrow e^{-2/3} - e^{-4/3}$.

2) The milage which a car owners get with a certain kind of radiar tyre is R.V having exponenti dist. with mean 40000 km. Find the prob that one of these tyres will last (i) atleast 20000 km.
ii) atmost 30,000 km

Mean $= 40000 = 1/\lambda$

$\lambda = \frac{1}{40000}$

$$\therefore f(x) = \begin{cases} \frac{1}{40000} e^{-x/40000}, & x \geq 0. \\ 0, & \text{otherwise}. \end{cases}$$

Let x = length of time.

i) $P(x \geq 2$

3) If $X \cong e^{\lambda}$ with $P(X \leq 1) = P(X > 1)$ find $var(X)$.

Exponential distribution.

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

Given,

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

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

$$2 \cdot P(X \leq 1) = 1.$$

$$P(X \leq 1) = 1/2.$$

$$\int_{-\infty}^{1} f(x)dx = 1/2$$

$$\lambda \int_{0}^{1} e^{-\lambda x} dx = 1/2$$

$$\lambda \left[\frac{e^{-\lambda x}}{-\lambda} \right]_0^1 = 1/2$$

$$\Rightarrow \quad -[e^{-\lambda} - 1] = 1/2.$$

$$-e^{-\lambda} + 1 = 1/2$$

$$1 - 1/2 = e^{-\lambda}.$$

$$e^{-\lambda} = 1/2$$

Apply 'log'

$$\log_e e^{-\lambda} = \log(1/2)$$

$$-\lambda \log_e e = \log_e 2^{-1}$$

$$-\lambda = -\log 2$$

$$\lambda = \log 2$$

$$Var(x) = \frac{1}{\lambda^2} \Rightarrow \frac{1}{(\log 2)^2}$$

4) If x is exp. distributed with par. λ, find the val of k, there exists $\frac{P(X > k)}{P(X \leq k)} = a$

Given

$$P(X > k) = a \cdot P(X \leq k).$$

$$P(X > k) = a[1 - P(X > k)]$$

Scanned by CamScanner

$P(x>k) = a - a\,P(x>k)$

$(1+a)\,P(x>k) = a.$

$P(x>k) = \frac{a}{1+a}$

$\int_k^\infty f(x)\,dx = \frac{a}{1+a}$

$\lambda \int_k^\infty e^{-\lambda x}\,dx = \frac{a}{1+a}$

$\lambda \left[\frac{e^{-\lambda x}}{-\lambda}\right]_k^\infty = \frac{a}{1+a}.$

$-\left[e^{-\infty} - e^{-\lambda k}\right] = \frac{a}{1+a}.$

$e^{-\lambda k} = \frac{a}{1+a}.$

'$\log_e$' Apply on B-S.

$\log_e e^{-\lambda k} = \log\left[\frac{a}{1+a}\right]$

$-\lambda k \log_e e = \log\left(\frac{a}{1+a}\right)$

$\boxed{k = \frac{-1}{\lambda}\log\left(\frac{a}{1+a}\right)}$, $\boxed{\therefore\ k = \frac{1}{\lambda}\log\left(\frac{1+a}{a}\right)}$

Continous Uniform Distribution:-

If X is a continous R.V if it follows prob. density function.

$$f(x) = \begin{cases} k, & \text{if } a \le x \le b \\ 0, & \text{otherwise.} \end{cases}$$

w.k.t

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\int_{a}^{b} k\,dx = 1$$

$$k[x]_a^b = 1$$

$$k[b-a] = 1$$

$$k = 1/b-a.$$

∴ Density function:-

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

Mean:-

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

$$\Rightarrow \int_a^b x \cdot \frac{1}{b-a}\,dx \Rightarrow \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b \Rightarrow \frac{1}{2(b-a)}(b^2-a^2)$$

$$E(x) = \frac{(b+a)(b-a)}{2}$$

$$E(x) = \text{mean}(\mu) = \frac{b+a}{2}.$$

Variance:-

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

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx.$$

$$E(x^2) = \frac{1}{b-a}\int_a^b x^2\,dx.$$

$$= \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b$$

$$E(X^2) = \frac{1}{3(b-a)}(b^3-a^3).$$

$$= \frac{(b-a)(b^2+ab+a^2)}{3(b-a)}$$

$$= \frac{b^2+ab+a^2}{3}.$$

$$V(X) = \frac{b^2+ab+a^2}{3} - \frac{(b+a)^2}{4}.$$

$$= \frac{1}{12}\left[4b^2+4ab+4a^2-3b^2-6ab-3a^2\right].$$

$$\Rightarrow \frac{1}{12}\left[b^2-2ab+a^2\right] \Rightarrow \frac{(b-a)^2}{12}.$$

$$V(X) = \frac{(b-a)^2}{12}.$$

~~Density~~ function:-

Distribution function:-

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx. \text{ if } a \leq x \leq b.$$

$$F(x) = \int_{a}^{x} f(x)\,dx.$$

$$= \int_{a}^{x} \frac{1}{b-a}\,dx \Rightarrow \frac{1}{b-a}[x]_a^x.$$

$$\Rightarrow \frac{[x-a]}{b-a}, \quad a \leq x \leq b.$$

$$F(x) = \begin{cases} 0 & , -\infty < x < a \\ \frac{x-a}{b-a} & , a \leq x \leq b \\ 1 & , b < x < \infty. \end{cases}$$

$F(-\infty) = 0$

$F(\infty) = 1.$

Note:-

if $x \in (-a, a)$

$$f(x) = \begin{cases} \frac{1}{2a} & , -a < x < a \\ 0, & \text{otherwise.} \end{cases}$$

1) A electric trains on a certain lines run every half an hour b/w midnight & 6:00 in the morning. what is the prob. that a man entering the station at random time during this period will have to ~~weigh~~ wait atleast 20 mins.

Let x = waiting time of a man.

$$f(x) = \begin{cases} \frac{1}{30-0}, & \text{if } 0 < x < 30 \\ 0, & \text{otherwise.} \end{cases}$$

prob of waiting time atleast 20 min.

$$P(x \geq 20) = \int_{20}^{30} f(x)dx \Rightarrow \int_{20}^{30} \frac{1}{30} \, dx \Rightarrow \frac{1}{30} [x]_{20}^{30}$$

$$\Rightarrow \frac{10}{30} \Rightarrow 1/3$$

2) Buses arrived to a specified bus stop at 15 mins intervals, starting at 7.am, ie., 7 am, 7:15 am, etc.. If a passenger arrives at same bustop at a random time which is uniformly distributed b/w 7:am and 7:30 am. Find the prob. that he waits (a) less than 5 mins (b) atleast 12 mins for a bus.

3) A passenger arrives at a local railway platform at 10 am knowing that the local train will arrive at some time uniformly distributed b/w 10 am and 10:30 am what is the prob. that (i) he will have to wait longer than 10 mins (ii) If at 10:15 am the train has not yet arrived, what is the prob. that he will have to wait atleast 10 additional mins.?

4) If x is uniformly distributed over (0, 10) find the prob. that (i) $x < 2$ (ii) $x > 8$ (iii) $3 < x < 9$.

5) If x is uniformly distributed with mean 2 & variance 4/3 then find $P(x < 0)$

Scanned by CamScanner

6) If x is uniformly distributed over $(-a, a)$ where $a > 0$. Find 'a'. Such that $P(X \geq 1) = 1/3$.

7) A R.V 'x' has uniform distribution $(-3, 3)$ find $|x+2| < 2$

$-2 < (x-2) < 2$ add 2

(i) $P(X < 2)$. (ii) $P(|x| < 2)$ iii) $P(|x-2| < 2)$

iv) Find 'k'. such that $P(x > k) = 1/3$

$P(0 < x < 4) = \int_0^4 f(x)dx$

$\Rightarrow \int_0^3 f(x)dx$.

2) let X = passenger arrives to bus stop at sometime to get bus.

$$f(x) = \begin{cases} \frac{1}{30-0}, & 0 < x < 30 \\ 0, & \text{otherwise} \end{cases}$$

i) prob. that he waits < 5 mins = prob. that he arrives bus stop at b/w 7:10 – 7:15 & 7:25 – 7:30.

$\Rightarrow P(10 < x < 15) + P(25 < x < 30)$.

$\Rightarrow \int_{10}^{15} f(x)\,dx + \int_{25}^{30} f(x)dx \Rightarrow \int_{10}^{15} \frac{1}{30} dx + \int_{25}^{30} \frac{1}{30} dx$

$\Rightarrow \frac{1}{30}[x]_{10}^{15} + \frac{1}{30}[x]_{25}^{30} \Rightarrow \frac{1}{30}[15-10] + \frac{1}{30}[30-25]$

$\Rightarrow 5/30 + 5/30 \Rightarrow \frac{10}{30} \Rightarrow 1/3$.

ii) prob. that he waits ≥ 12 mins = prob. that he arrives bus stop b/w 7:00 – 7:03 & 7:15 – 7:18.

$\Rightarrow P(0 < x < 3) + P(15 < x < 18)$.

$\Rightarrow \int_0^3 f(x)\,dx + \int_{15}^{18} f(x)dx \Rightarrow \int_0^3 \frac{1}{30} dx + \int_{15}^{18} \frac{1}{30} dx$

$\Rightarrow \frac{1}{30}[x]_0^3 + \frac{1}{30}[x]_{15}^{18} \Rightarrow 3/30 + 3/30 \Rightarrow 6/30 \Rightarrow 1/5$.

3) Let X = passenger arrival time to railway station at some time.

$$f(x) = \begin{cases} \frac{1}{30-0}, & 0 < x < 30 \\ 0, & \text{otherwise.} \end{cases}$$

i) prob. that he waits ≥ 10 mins = prob. that he arrives railway station b/w 10:10 – 10:30

$\Rightarrow P(10 \le X < 30)$

$\Rightarrow \int_{10}^{30} f(x)\,dx \Rightarrow \int_{10}^{30} \frac{1}{30}\,dx \Rightarrow \frac{1}{30}[30-10] \Rightarrow 2/3.$

ii) To find prob. that the passenger has to wait atleast 10 additional mins given that the train has not arrived at 10:15 am.

P(he has to wait atleast 10 additional mins given that he has already waited 15 mins) $= P(X > (15+10) / X > 15)$.

$\Rightarrow P(X>25 / X>15) = \frac{P(X>25 \cap X>15)}{P(X>15)} \Rightarrow \frac{P(X>25)}{P(X>15)}$

$P(X>25) = \frac{\int_{25}^{30} f(x)dx}{\int_{15}^{30} f(x)dx} \Rightarrow \frac{\frac{1}{30}[x]_{25}^{30}}{\frac{1}{30}[x]_{15}^{30}} \Rightarrow \frac{30-25}{30-15} \Rightarrow \frac{5}{15} \Rightarrow 1/3.$

4> Given $(0,10) = (a,b)$.

$f(x) = \begin{cases} \frac{1}{10} & , 0 < x < 10 \\ 0 & , \text{otherwise} \end{cases}$

i) $P(x<2) = \int_0^2 f(x)dx \Rightarrow \frac{1}{10}\int_0^2 1\,dx \Rightarrow \frac{1}{10}[x]_0^2 \Rightarrow 2/10 = 1/5.$

ii) $P(x>8) = \int_8^{10} f(x)dx \Rightarrow \frac{1}{10}\int_8^{10} 1.dx \Rightarrow \frac{1}{10}[10-8] = 2/10 = 1/5.$

iii) $P(3<x<9) \Rightarrow \int_3^9 f(x)dx \Rightarrow \frac{1}{10}[9-3] \Rightarrow 6/10 \Rightarrow 3/5.$

5) Given

Mean = 1

$\frac{a+b}{2} = 1$

$a+b = 2$

var = 4/3

$\frac{(b-a)^2}{12} = 4/3$

$(b-a)^2 = 16$

$b-a = \pm 4$

$b-a = 4$

$\Rightarrow b+a = 2$

$b-a = 4$

$2b = 6$

$b = 3$

$a = 2-3$

$a = -1$

$(-1, 3)$

$\Rightarrow b+a = 2$

$b-a = -4$

$2b = -2$

$b = -1$

$a = 2+1$

$a = 3.$

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

$\therefore$ Interval $\Rightarrow (-1, 3)$

$$f(x) = \begin{cases} \frac{1}{3+1} & , -1 \le x \le 3 \\ 0 & , \text{otherwise.} \end{cases}$$

$$P(X<0) = \int_{-1}^{0} f(x)\,dx \Rightarrow \int_{-1}^{0} 1/4\,dx \Rightarrow 1/4[x]_{-1}^{0} \Rightarrow 1/4 .$$

6> Interval $\Rightarrow (-a, a)$; $a>0$

$$f(x) = \begin{cases} \frac{1}{a-(-a)} & , -a<x<a \\ 0 & , \text{otherwise} \end{cases} \Rightarrow \begin{cases} \frac{1}{2a} & , -a<x<a \\ 0 & , \text{otherwise.} \end{cases}$$

Given, $P(X \ge 1) = 1/3$

$$\int_{1}^{\infty} f(x)\,dx = 1/3$$

$$\int_{1}^{a} f(x)\,dx = 1/3 \Rightarrow \int_{1}^{a} \frac{1}{2a}\,dx = 1/3$$

$$\frac{1}{2a}[x]_1^a \Rightarrow 1/3$$

$$\frac{a-1}{2a} = 1/3$$

$$3a - 3 = 2a$$

$$\boxed{\therefore a = 3}$$

Gamma Distribution :-

The continous R.V 'x' is said to follow gamma distribution if the density function, .

$$f(x) = \begin{cases} \frac{\lambda^{\alpha}}{\Gamma\alpha} \cdot e^{-\lambda x} \cdot x^{\alpha-1} , & \alpha ; \lambda > 0, \; 0<x<\infty . \\ 0 & , \text{otherwise.} \end{cases}$$

* $\Gamma(n) = \int_{0}^{\infty} e^{-x} \cdot x^{n-1}\,dx$

$\Gamma n = (n-1)\Gamma(n-1)$. if n is +ve function.

$\Gamma n = (n-1)!$ if n is +ve integral.

$\Gamma 1 = 1$

$\Gamma 1/2 = \sqrt{\pi}$

Mean :-

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

$$\Rightarrow \int_0^{\infty} x \frac{\lambda^{\alpha}}{\Gamma(\alpha)} e^{-\lambda x} \cdot x^{\alpha-1}\,dx$$

$$\Rightarrow \frac{\lambda^{\alpha}}{\Gamma\alpha} \int_0^{\infty} e^{-\lambda x} \cdot x^{\alpha}\,dx$$

Let $\lambda x = t$

$x = t/\lambda$, $dx = \frac{dt}{\lambda}$

$$E(x) = \frac{\lambda^{\alpha}}{\Gamma\alpha} \int_0^{\infty} e^{-t} \left(\frac{t}{\lambda}\right)^{\alpha} \cdot \frac{dt}{\lambda}$$

$$E(x) = \frac{1}{\lambda\Gamma\alpha} \int_0^{\infty} e^{-t} \cdot t^{\alpha}\,dt$$

$$= \frac{1}{\lambda\Gamma\alpha} \int_0^{\infty} e^{-t} \cdot t^{(\alpha+1)-1}\,dt$$

$$= \frac{1}{\lambda\Gamma\alpha} \Gamma\alpha+1 \Rightarrow \frac{1}{\lambda\Gamma\alpha} \alpha \cdot \Gamma\alpha$$

$$\boxed{\therefore \text{Mean} = \frac{\alpha}{\lambda}}$$

if $1/\lambda = \beta$, $\boxed{\therefore \text{Mean} = \alpha\beta}$

Variance :-

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

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$$

$$E(x^2) = \frac{\lambda^{\alpha}}{\Gamma\alpha} \int_0^{\infty} x^2 e^{-\lambda x} \cdot x^{\alpha-1}\,dx$$

$$E(x^2) = \frac{\lambda^{\alpha}}{\Gamma\alpha} \int_0^{\infty} e^{-\lambda x} \cdot x^{\alpha+1}\,dx$$

Let $\lambda x = t$

$x = t/\lambda$ $\Rightarrow dx = \frac{dt}{\lambda}$

$$E(x^2) = \frac{\lambda^{\alpha}}{\Gamma\alpha} \int_0^{\infty} e^{-t} \left(\frac{t}{\lambda}\right)^{\alpha+1} \cdot \frac{dt}{\lambda}$$

$$E(x^2) = \frac{\cancel{\lambda^{\alpha}}}{\Gamma\alpha} \frac{1}{\cancel{\lambda^{\alpha}} \cdot \lambda^2} \int_0^{\infty} e^{-t} \cdot t^{\alpha+1}\,dt$$

$$E(x^2) = \frac{1}{\lambda^2\Gamma\alpha} \Gamma(\alpha+2) \quad \frac{1}{\lambda^2\Gamma\alpha} \int_0^{\infty} e^{-t} \cdot t^{(\alpha+2)-1}\,dt$$

Scanned by CamScanner

$$\Rightarrow \frac{1}{\lambda^2 \Gamma\alpha} \Gamma(\alpha+2)$$

$$\Rightarrow \frac{1}{\lambda^2 \Gamma\alpha} (\alpha+1) \Gamma(\alpha+1)$$

$$\Rightarrow \frac{1}{\lambda^2 \Gamma\alpha} (\alpha+1) \alpha \Gamma\alpha$$

$$E(x^2) = \frac{\alpha^2+\alpha}{\lambda^2}$$

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

$$= \frac{\alpha^2+\alpha}{\lambda^2} - \frac{\alpha^2}{\lambda}$$

$$\boxed{V(x) = \frac{\alpha}{\lambda^2}}$$

if $1/\lambda = \beta$ then $\boxed{V(x) = \alpha\beta^2}$

1) In a certain city, the dialy consumption of electric power in millions of kwh can be treated as a R.V. having Gamma distribution with parameters $\lambda = 1/2, \alpha = 3$ If the power plant of this city daily capacity of 12 mkwh, what is the prob. that this power supply will be inadequate on any given day?

Given $\lambda = 1/2$, $\alpha = 3$

x = power consumption in a city.

$$f(x) = \frac{\lambda^\alpha}{\Gamma\alpha} e^{-\lambda x} x^{\alpha-1}$$

$$f(x) = \left(\frac{1}{2}\right)^3 \cdot \frac{1}{\Gamma 3} e^{-x/2} \cdot x^2, \quad 0 < x < \infty$$

$$= \frac{1}{2^3} \cdot \frac{1}{2!} x^2 e^{-x/2}, \quad 0 < x < \infty$$

$$f(x) \Rightarrow \frac{x^2 e^{-x/2}}{16}, \quad 0 < x < \infty$$

$$P(x > 12) = \int_{12}^{\infty} f(x)\,dx \Rightarrow \int_{12}^{\infty} \frac{x^2 e^{-x/2}}{16}\,dx.$$

$$\Rightarrow \frac{1}{16}\left[x^2\left(\frac{e^{-x/2}}{-1/2}\right) - 2x\left(\frac{e^{-x/2}}{1/4}\right) + 2\left(\frac{e^{-x/2}}{-1/8}\right)\right]_{12}^{\infty}$$

$\Rightarrow \frac{1}{16}\{(0+0+0) - e^{-6}(-2(12)^2 - 4(24) + (-16))\}$

$\Rightarrow \frac{e^{-6}}{16}[288 + 96 + 16] \Rightarrow \frac{e^{-6}}{16}[400] \Rightarrow 25e^{-6}.$

2) The daily consumption of milk in a city, in excess of 20,000 litres is approximately distributed as a gamma variate with parameters $\alpha = 2$, $\lambda = \frac{1}{10,000}$. The city has daily stock of 30,000 litres, what is the prob. that the stock is insufficient on particular day?

$\lambda = \frac{1}{10,000}$, $\alpha = 2$.

X = excess of milk consumption. / Y = daily consumption.

$Y = X + 20,000$ then $\Rightarrow X = Y - 20,000$.

$f(x) = \frac{1}{(10,000)^2}\frac{1}{\lceil 2}e^{-x/10,000}\cdot x$

$f(x) = \frac{x}{(10,000)^2}\cdot e^{-x/10,000}$, $0 < x < \infty$.

$P(Y > 30,000) = P(X + 20000 > 30000)$

$= P(X > 10,000)$

$= \int_{10,000}^{\infty} f(x)\,dx \Rightarrow \int_{10,000}^{\infty}\frac{x}{(10,000)^2}\cdot e^{-x/10000}\,dx.$

$\Rightarrow \frac{1}{(10,000)^2}\int_{10,000}^{\infty} x\cdot e^{-x/10000}\,dx \Rightarrow \frac{1}{(10,000)^2}\left[x\left(\frac{e^{-x/10000}}{-1/10000}\right) - 1\left(\frac{e^{-x/10000}}{1/10000^2}\right)\right]_{10,000}^{\infty}$

$\Rightarrow \frac{(10,000)^2}{(10,000)^2}\left[10000xe^{-x/10000} - e^{-x/10000}\right]_{10,000}^{\infty}$

$\Rightarrow \{(0+0) - ((-10000)^2e^{-1} - e^{-1})\}$

$\Rightarrow \frac{1}{e}[(10000)^2+1] \Rightarrow$

Sampling Distributions :-

Population :- The aggregate or totality of statistical data corresponding to observations whether it is either finite or infinite is called population.

Ex :- i) population of heights of indians.

ii) No. of students in a college those are classified according to blood group.

Size of population :- The no. of observations in the population is known as size of population. It is denoted by 'N'.

Sample :- A small portion of population which gives population characteristics is known as sample (or) The subset of population is known as sample.

Size of sample :- The no. of observations in sample is known as ~~Sam~~ size of sample. It is denoted by 'n'.

Large sample :- If the size of sample $n \geq 30$, then sample is called large sample.

Small sample :- If the size of sample $n < 30$, then the sample is called small sample.

Sampling :- The process of selection of sample from the population is known as sampling.

* There are 2 types of sampling distributions

1) probability sampling distributions :-

i) Random sampling distribution

ii) Stratified " "

iii) Systematic " "

2) Non-probability sampling distributions :-

i) purposive sampling distribution

ii) sequential " "

Note :-

i) No. of samples with replacement from population $= N^n$

ii) No. of samples without replacement from population $= {}^NC_n$

iii) sampling with replacement considered as infinite population.

iv) sample without replacement finite population.

Symbols

population

i) population mean $= \mu$

$$\mu = \frac{1}{N}\sum_{1} x_i$$

ii) pop. variance

$$\sigma^2 = \frac{1}{N}\sum_{1}(x_i-\mu)^2$$

iii) N = size of population

sample.

i) Sample mean $= \bar{x}$

$$\bar{x} = \frac{1}{n}\sum_{1} x_i$$

ii) sample variance

$$s^2 = \frac{1}{n-1}\sum_{1}(x_i-\bar{x})^2$$

iii) n = size of sample.

Parameters :-

The statistical measurements of all observations of populations is known as parameters.

Ex:- μ, σ^2.

Statistics :-

The statistical measurements of all units selected in a sample is known as statistics.

Ex:- $\bar{x}$, s^2.

Sampling distributions of mean :-

Let $x_1, x_2, x_3, \ldots x_n$ be 'n' random samples with size 'n' drawn from population of size 'N' with mean μ & variance of, let $\bar{x}$ be mean of sample

1) Infinite population :- (Sampling with replacement) :-

i) mean of popⁿ $= \mu = \frac{1}{N}\sum x_i$

ii) variance of pop $= \sigma^2 = \frac{1}{N}\sum (x_i - \mu)^2$

iii) Mean of sampling distribution of means $= \mu_{\bar{x}}$

$$\boxed{\mu_{\bar{x}} = \mu}$$

iv) variance of sampling distribution of means $= \sigma_{\bar{x}}^2$

$$\sigma_{\bar{x}}^2 = \frac{1}{\text{total samples}}\left[\sum(\bar{x}_i - \mu_{\bar{x}})^2\right] \quad \text{(or)}$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

2) Finite population :- (sampling without Replacement).

i) Mean of pop. $= \mu$

ii) variance of pop $= \sigma^2$

iii) Mean of sampling dist. of means $= \mu_{\bar{x}}$

$$\mu_{\bar{x}} = \mu.$$

iv) Variance of sampling distr. of means $= \sigma_{\bar{x}}^2$

$$\sigma_{\bar{x}}^2 = \frac{1}{\text{total}}\sum(\bar{x}_i - \mu_{\bar{x}})^2 \quad \text{(or)} \quad \sigma_{\bar{x}}^2 = \left(\frac{N-n}{N-1}\right)\frac{\sigma^2}{n}$$

Finite population correction factor :-

$$\frac{N-n}{N-1}.$$

Central limit theorem :-

If $\bar{x}$ is ~~central limit~~ sample mean ~~of~~ with size n drawn from a population having mean μ & Standard deviation σ, then the sampling distribution of mean follows normal distribution by $\boxed{z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}}$

Standard error of sample mean :-

$$S.E.(\bar{x}) = \frac{\sigma}{\sqrt{n}}$$

Note:-

1) For infinite population, the sampling distribution of sum's of mean's has the mean $\mu_{\bar{x}_1+\bar{x}_2}$ and S.D $\sigma_{\bar{x}_1+\bar{x}_2}$ is defined by

$$\mu_{\bar{x}_1+\bar{x}_2} = \mu_{\bar{x}_1} + \mu_{\bar{x}_2}$$

$$\sigma_{\bar{x}_1+\bar{x}_2} = \sqrt{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

2) Sampling distribution of differences of means has the mean $\mu_{\bar{x}_1-\bar{x}_2}$ & S.D $\sigma_{\bar{x}_1-\bar{x}_2}$ are defined by

$$\mu_{\bar{x}_1-\bar{x}_2} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2}$$

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

1) A population consists of 5 numbers 2, 3, 6, 8 & 11. Consider all possible samples of size '2' which can be drawn from this population. Then find (i) mean of population (ii) Standard deviation of population. (iii) mean of sampling distribution of means (iv) S.D of sampling distribution of mean.

Given population is 2, 3, 6, 8 & 11

population size (N) = 5

Sample size (n) = 2

i) mean of population (μ) :-

$$\mu = \frac{2+3+6+8+11}{5} \Rightarrow \frac{30}{5} \Rightarrow 6.$$

ii) variance of population (σ^2) :-

$$\sigma^2 = \frac{1}{N} \sum (x_i-\mu)^2$$

$$\Rightarrow \frac{1}{5}\left[(2-6)^2 + (3-6)^2 + (6-6)^2 + (8-6)^2 + (11-6)^2\right]$$

$$\Rightarrow \frac{1}{5}[16+9+0+4+25]$$

$$\Rightarrow \frac{1}{5}(54)$$

$$\Rightarrow 10.8$$

$$\sigma = \sqrt{10.8} \Rightarrow 3.$$

iii) No. of samples with replacements $= N^2 = 5^2 = 25$

Samples $= \{ (2,2), (2,3), (2,6), (2,8), (2,11)$
$(3,2), (3,3), (3,6), (3,8), (3,11)$
$(6,2), (6,3), (6,6), (6,8), (6,11)$
$(8,2), (8,3), (8,6), (8,8), (8,11)$
$(11,2), (11,3), (11,6), (11,8), (11,11) \}$

All samples are $= \{ 2, 2.5, 4, 5, 6.5,$
$2.5, 3, 4.5, 5.5, 7,$
$4, 4.5, 6, 7, 8.5,$
$5, 5.5, 7, 8, 9.5,$
$6.5, 7, 8.5, 9.5, 11 \}$

Mean of sampling distributions of means.

$$\mu_{\bar{x}} = \frac{2+2.5+4+5+ \cdots +9.5+11}{25}$$

$$\mu_{\bar{x}} = 6.$$

iv) S.D of sampling distribution.

$$\sigma^2_{\bar{x}} = \frac{1}{total} \sum (\bar{x}_i - \mu_{\bar{x}})^2$$

$$= \frac{1}{25} [(2-6)^2 + (2.5-6)^2 + \cdots + (11-6)^2].$$

$$\Rightarrow 5.4.$$

$$\sigma_{\bar{x}} = \sqrt{5.4} \Rightarrow 2.32.$$

(or)

$$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n} = \frac{10.8}{2} \Rightarrow 5.4. \quad (\text{w.k.t})$$

2) Solve above problem for without replacement:-

Given population is 2, 3, 6, 8 & 11.

i) (ii) questions are same as above

iii) No. of samples without replacement $= {}^N C_n = {}^5 C_2 = 10$

Samples $= \{ (2,3), (2,6), (2,8), (2,11)$
$(3,6), (3,8), (3,11),$
$(6,8), (6,11)$
$(8,11) \}$

All Samples means = $\{2.5, 4, 5, 6.5, 4.5, 5.5, 7, 7, 8.5, 9.5\}$.

Mean of Sampling distribution of means = :

$$\mu_{\bar{x}} = \frac{2.5+4+5+\cdots+8.5+9.5}{10} \Rightarrow 6.$$

iv) $\sigma_{\bar{x}}^2 = \frac{1}{total} \sum (\bar{x_i} - \mu_{\bar{x}})^2$

$\sigma_{\bar{x}}^2 = \frac{1}{10}[(2.5-6)^2 + (4-6)^2 + \cdots + (9.5-6)^2]$.

$\sigma_{\bar{x}}^2 = 4.05$

$\sigma_{\bar{x}} = \sqrt{4.05} \Rightarrow 2.01$.

(or)

$\sigma_{\bar{x}}^2 = \left(\frac{N-n}{N-1}\right)\frac{\sigma^2}{n} \Rightarrow \left(\frac{5-2}{5-1}\right)\frac{10.8}{2} = \left(\frac{3}{4}\right)\left(\frac{10.8}{2}\right) = 4.05$

H.W

2) A population is 5, 10, 14, 18, 13, 24. consider all possible samples of size '2' which can be drawn without replacement from this population. Find (i) Mean of population (ii) S.D of population iii) Mean of sampling dist. of means iv) S.D of sampling dist. of means.

3) If the population 3, 6, 9, 15 & 27. (i) Find all possible samples of size '3' that can be taken without replacement from the finite population. (ii) mean of sampling dist. of means (iii) S.D of all means

Given population is 3, 6, 9, 15, 27

Population size (N) = 5

Sample size (n) = 3

Mean of population (μ) :-

$$\mu = \frac{3+6+9+15+27}{5} = \frac{60}{5} \Rightarrow 12$$

variance of population.

$\sigma^2 = \frac{1}{N}\sum (x_i - \mu)^2$

$= \frac{1}{5}[(3-12)^2 + (6-12)^2 + (15-12)^2 + (27-12)^2]$

$\Rightarrow \frac{1}{5}[81+36+9+9+25] \Rightarrow \frac{1}{5}(360) = 72$

$\sigma^2 = 72$

$\sigma = \sqrt{72} = 8.48$.

Good

8) A Random Sample of size 64 is taken from a normal population with mean $\mu = 51.4$ and $\sigma = 6.8$. What is the probability that the mean of sample will

i) exceed 52.9 ii) fall b/w 50.5 & 52.3

iii) less than 50.6.

~~9) The mean height of students in a clg 155 cm~~

9) A Random sample of size 100 is taken from infinite population having mean $\mu = 76$ $\sigma^2 = 256$. What is the probability that $\bar{x}$ lies b/w 75 & 78.

10) The mean breaking strength of copper wire is 575 lbs with S.D 8.3 lbs. How large a sample must be used inorder that there will be one chance in 100 that mean v. strength of sample is less than 572 lbs

11) 3 masses are measured as 62.34, 20.48, 35.97 kgs S.D = 0.54, 0.21, 0.46 kgs. find mean and S.D of sum of masses.

8)

Mean = 51.4. Sample size (n) = 64

S.D(σ) = 6.8

i) To find $P(\bar{x} > 52.9)$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \Rightarrow$$

If $\bar{x}$ 52.9 $\Rightarrow z = \dfrac{52.9 - 51.4}{6.8/\sqrt{64}} \Rightarrow 1.87$

$P(\bar{x} > 52.9) = P(z > 1.87)$

$= 0.5 - P(0 < z < 1.87)$

$= 0.5 - A(1.87)$

$= 0.5 - 0.4693 \Rightarrow 0.0307$

ii) $P(50.5 < \bar{x} < 52.3)$

if $\bar{x} = 50.5 \Rightarrow z = \dfrac{50.5 - 51.4}{6.8/\sqrt{64}} \Rightarrow \overset{-1.05}{\cancel{-1.125}} \Rightarrow \cancel{\cong -1.13}$

if $\bar{x} = 52.3 \Rightarrow z = \dfrac{52.3 - 51.4}{6.8/\sqrt{64}} \Rightarrow \overset{-1.06}{\cancel{1.125}} \cong \cancel{1.13}.$

$P(-1.06 < z < 1.06) = P(-1.06 < z < 0) + P(0 < z < 1.06)$

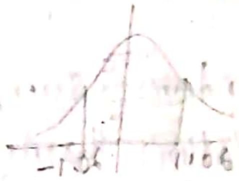
$\Rightarrow P(0 < z < 1.06) + A(1.06)$

$\Rightarrow A(1.06) + A(1.06)$

$\Rightarrow 2A(1.06)$

$\Rightarrow 2(0.3554)$

$\Rightarrow 0.7108.$



iii) $P(\bar{x} < 50.6)$

if $\bar{x} = 50.6 \Rightarrow z = \dfrac{50.6 - 51.4}{6.8/\sqrt{64}} = -0.94$

$P(\bar{x} < 50.6) = P(z < -0.94)$

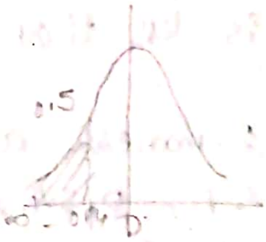
$\Rightarrow 0.5 - P(-0.94 < z < 0)$

$\Rightarrow 0.5 - P(0 < z < 0.94)$

$\Rightarrow 0.5 - A(0.94)$

$\Rightarrow 0.5 - 0.3264$

$\Rightarrow 0.1736.$



9) Mean $= \mu = 76.$

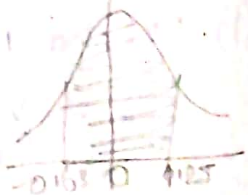
$S.D(\sigma) = \sqrt{256} = 16.$

sample size $(n) = 100.$

i) $P(75 < \bar{x} < 78)$

if $\bar{x} = 75 \Rightarrow z = \dfrac{75 - 76}{16/\sqrt{100}} = -0.63$

if $\bar{x} = 78 \Rightarrow z = \dfrac{78 - 76}{16/\sqrt{100}} = 1.25.$



$P(-0.63 < z < 1.25) = P(-0.63 < z < 0) + P(0 < z < 1.25)$

$= A(-0.63) + A(1.25)$

$= 0.2357 + 0.3944$

$\Rightarrow$

Scanned by CamScanner

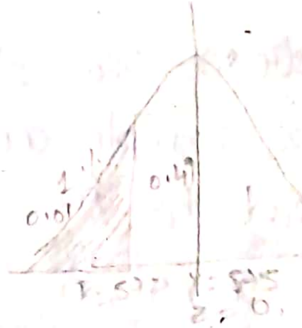
10) Mean(μ) = 575 lbs

S.D(σ^2) = 8.3 lbs

if $\bar{x}$ = 572 lbs $\Rightarrow z = \frac{x-\mu}{\sigma/\sqrt{n}} \Rightarrow \frac{572-575}{8.3/\sqrt{n}} \Rightarrow \frac{-3\sqrt{n}}{8.3}$

$P(\bar{x} < 572) = \frac{1}{100} = 1\% \Rightarrow 0.01$.

from diagram.



$P(-z_1 < z < 0) = 0.49$

$A(-z_1) = 0.49$

$z_1 = -2.33$

$\frac{-3\sqrt{n}}{8.3} = -2.33$

$\sqrt{n} = \frac{2.33 \times 8.3}{3}$

$n = \left(\frac{2.33 \times 8.3}{3}\right)^2$

$n = 42$.

11) Assume masses,

$\mu_{\bar{x}_1} = 62.34$

$\mu_{\bar{x}_2} = 20.48$

$\mu_{\bar{x}_3} = 35.97$.

Mean of sum of masses $\Rightarrow \mu_{\bar{x}_1} + \mu_{\bar{x}_2} + \mu_{\bar{x}_3} = \mu_{\bar{x}_1+\bar{x}_2+\bar{x}_3}$

$\Rightarrow 62.34 + 20.48 + 35.97$

$\Rightarrow$

S.D of sum of masses $\sigma_{\bar{x}_1+\bar{x}_2+\bar{x}_3} = \sqrt{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2} + \sigma^2_{\bar{x}_3}}$.

$= \sqrt{(62.34)^2 + (20.48)^2 + (35.97)^2}$

$=$

UNIT – IV

Estimation & Tests of Hypotheses

Short Answer questions

1) Define Null Hypothesis and Alternative Hypothesis

Sol:- **Null Hypothesis** :- A definite statement about the population parameter for applying the test of significance is called null hypothesis, which is usually a hypothesis of no difference and is denoted by H_0

Alternative Hypothesis :- Any hypothesis which is complementary to the null hypothesis is called an alternative hypothesis and is denoted by H_1

2) Define Type I error and Type II error

Type I error :- If we reject the null hypothesis when it is true, and it is also known as producer risk

$$P[\text{reject } H_0 \text{ when it is true}] = \alpha$$

Type II error :- If we accept the null hypothesis when it is wrong, and it is also known as consumer risk.

$$P[\text{accept } H_0 \text{ when it is wrong}] = \beta$$

(3) if a Sample number is '500' and The S.D is '15' find maximum error with 95% Confidence.

Sol:-

Maximum error $E = Z_{\alpha/2} \sigma/\sqrt{n}$

$Z_{\alpha/2} = 1.96$ at 95%.

$\sigma = 15$

$n = 500$

$E = 1.96 \dfrac{15}{\sqrt{500}}$

$E = 1.31$

(4) The mean & S.D of population are 11795 & 14504 respectively if n=50. find 95% Confidence Interval for The mean

Sol:-

$$\text{Interval} = \left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \; \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

$\bar{x} = 11795$ $\quad \sigma = 14,504$

$$= \left(11795 - (1.96)\frac{14054}{\sqrt{50}}, \; 11795 + (1.96)\frac{14054}{\sqrt{50}}\right)$$

$$= (7899.42, \; 15690.57)$$

(5) if we can assert with 95% that The maximum error is 0.05 and $p=0.2$ find Sample Size

Sol:-

Max error $E = Z_{\alpha/2}\sqrt{\dfrac{PQ}{n}}$

$p = 0.2$ $\quad Q = 1-P$
$= 1-0.2$
$= 0.8$

$$0.05 = 1.96\sqrt{\frac{(0.2)(0.8)}{n}}$$

$$n = \frac{(0.2)(0.8) \times (1.96)^2}{(0.05)^2} = 246$$

(6) A Random Sample of size '100' has a S.D of '5'. What can u say about the max error with 95% confidence

Sol: Given That $n=100$ $\sigma=5$

$$Z_{\alpha/2}=1.96 \quad (95\% \text{ Confidence})$$

Max error $E = Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$

$$= 1.96 \frac{5}{\sqrt{100}} = 0.98$$

(1) a) Construct 95% Confidence Interval for True proportion of Computer literates if ~~47~~ 47 out of 150 persons from rural areas are Computer literates

proportion $p = \dfrac{x_1}{n_1} = \dfrac{47}{150} = 0.313$ $\quad Q = 1-P$

$= 1-0.313$

$= 0.687$

Confidence Interval

$$\left(p - Z_{\alpha/2}\sqrt{\frac{PQ}{n}}\ ,\ p + Z_{\alpha/2}\sqrt{\frac{PQ}{n}}\right)$$

$$\left(0.313-(1.96)\sqrt{\frac{(0.313)(0.687)}{150}}\ ,\ 0.313+1.96\sqrt{\frac{(0.313)(0.687)}{150}}\right)$$

$$(0.313 - 0.0742\ ,\ 0.313+0.0742)$$

$$(0.2388\ ,\ 0.3872)$$

b) A Sample of size 9 was taken from a population gave $s^2 = 10.9$, $\bar{x} = 15.8$ obtain 99% Confidence Interval for μ

$$\left(\bar{x} \pm z_{\alpha/2}\, s/\sqrt{n}\right)$$

$$\left(15.8 - 2.58\left(3.3/\sqrt{9}\right),\ 15.8 + 2.58\left(3.3/\sqrt{9}\right)\right)$$

$$(12.96,\ 18.6)$$

Critical Value (significant values):- The value of test statistics separates a rejection region and the acceptance region is called the Critical value.

Critical value of z	level of Significance (α)		
	1%	5%	10%
Two tailed	2.58	1.96	1.64
Right tailed	2.33	1.64	1.28
Left tailed	-2.33	-1.64.	-1.28

long Answer questions

2) The mean height of 90 Students in a Class is 180cm test 90% level whether the Sample has been drawn from a population mean is 170 Cms and standard deviation '35'.

Sol:- Given data. Sample mean $\bar{x} = 180$

Sample Size $n = 90$

population mean $\mu = 170$

population S.D $\sigma = 35$

Null Hypothesis: Ho :- The Sample has been drawn from a population $(\bar{x} = \mu)$

Alternative Hypothesis H_1: The Sample has not drawn from a population $(\bar{x} \neq \mu)$ (Two Tailed)

Test Statistic: $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$

$= \dfrac{180 - 170}{35/\sqrt{90}} = \dfrac{10 \times 9.48}{35} = 2.71$

$z_{cal} =$

level of Significance: $\alpha = 10\%$

$z_{tab} = 1.64$

Conclusion: $|z_{cal}| > z_{tab}$ $(|2.71| > 1.64)$

So Alternative Hypothesis accepted

The Sample is not drawn from Same population

③ The means of two large samples of sizes 1000 & 2000 member are 67.5 Inches & 68.0 Inches Can the sample be regarded as drawn from the same population of S.D 2.5 Inches.

Sol:- Given data

first sample mean $\bar{x}_1 = 67.5$

second sample mean $\bar{x}_2 = 68$

pop S.D $= \sigma_1 = \sigma_2 = \sigma = 2.5$

$n_1 = 1000 \quad n_2 = 2000$

Null Hypothesis H_0: Two samples drawn from the same population i.e ($\mu_1 = \mu_2$)

Alternative Hypothesis H_1: Two samples not drawn from the same population i.e ($\mu_1 \neq \mu_2$)

Test statistic $z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma^2}{n_1} + \dfrac{\sigma^2}{n_2}}} = \dfrac{67.5 - 68}{\sqrt{\dfrac{(2.5)^2}{1000} + \dfrac{(2.5)^2}{2000}}}$

$= -5.16$

level of significance $\alpha = 5\% \quad z_\alpha = 1.96$

Conclusion: $|z|_{cal} > z_{tab} \quad (|-5.16| > 1.96)$

Alternative hypothesis accepted

The samples are drawn from the same population

4) A manufacturer Claimed that at least 98% of The Steel Pipes which he Supplied to a factory Conformed to Specifications. ~~test this Claim at a Significance level of 0.05~~. An examination

~~population proportion $P = \frac{98}{100} = 0.98$~~

of 500 pieces of pipes revealed That '30' were defective. test this Claim at a Significance level of 0.05

Sol:-

population proportion $P = \frac{98}{100} = 0.98$

500 pipes '30' are defective So non defective

pipes $= 500 - 30 = 470$

Sample proportion of non defective pipes $p = \frac{x_1}{n_1} = \frac{470}{500}$

$= 0.94$

$n_1 = 500$

Null Hypothesis H_0:- There is no diff b/w

Sample proportion with population proportion

i.e $(p = P)$

Alternative Hypothesis H_1 There is diff b/w

Sample proportion with population proportion

i.e $(p \neq P)$

Test Statistic:- $z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$

$P = 0.98$ $\quad Q = 1 - P$

$= 1 - 0.98$

$= 0.02$

$$z_{cal} = \frac{0.94 - 0.98}{\sqrt{\frac{(0.98)(0.02)}{500}}} = -2.02$$

leve of Significance: $\alpha = 0.05$

$= 5\%$

$z_{tab} = 1.96$

Conclusion:- $|z_{cal}| > z_{tab}$ $\quad (|-2.02| > 1.96)$

Alternative Hypothesis accepted

There is diff b/w Sample proportion with pop proportion.

(5) In a Certain City 125 men in a Sample of 500 were found to be Smokers. In another City The number of Smokers was 375 in a random Sample of 1000. Does this indicate that there is a greater population smokers in the Second City than the first City?

Sol:-

Smokers proportion In first City $p_1 = \frac{x_1}{n_1} = \frac{125}{500} = 0.25$

Smokers proportion In second City $p_2 = \frac{x_2}{n_2} = \frac{375}{1000} = 0.375$

Null Hypothesis H_0: There is no diff b/w Two City Smokers proportion $(p_1 = p_2)$

Alternative Hypothesis H_1:- There is a greater population Smokers in the Second City than the first city $(p_2 > p_1)$ (Right tailed)

Test Statistic

$$z = \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$p = \frac{x_1 + x_2}{n_1 + n_2} = \frac{125 + 375}{500 + 1000} = \frac{500}{1500} = 0.\dot{3}3$$

$$q = 1 - p$$
$$= 1 - 0.33$$
$$= 0.67$$

$$Z = \frac{(0.25) - (0.375)}{\sqrt{(0.33)(0.67)\left(\frac{1}{500} + \frac{1}{1000}\right)}} = 1.506$$

Level of Significance: $\alpha = 5\%$ (Right tailed)

$$Z_{tab} = 1.64$$

Conclusion: $|Z_{cal}| < Z_{tab}$ $(|1.506| < 1.64)$

Null Hypothesis accepted

There is a greater population Smokers in the Second City then the first City

(6) The owner of a machine shop must decide which of two snack vending machines to install in this shop. if each is tested '250' times the first machine fails to work 13 times and second machine fails to work '7' times. test at 0.05 level of significance whether the difference b/w the corresponding sample proportions is significant

Sol:-

first sample size $n_1 = 250$

Second sample size $n_2 = 250$

first sample proportion to fails machine $p_1 = \frac{x_1}{n_1} = \frac{13}{250} = 0.052$

second sample proportion $p_2 = \frac{x_2}{n_2} = \frac{7}{250} = 0.028$

$p_1 = 0.052 \quad p_2 = 0.028$

Null Hypothesis H_0:- $p_1 = p_2$

Alternative Hypothesis: H_1:- $p_1 \neq p_2$

Test Statistic $$z = \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$p = \frac{x_1 + x_2}{n_1 + n_2} = \frac{13+7}{250+250} = \frac{20}{500} = 0.04$$

$$q = 1 - 0.04 = 0.96$$

$$z_0 = \frac{0.052 - 0.028}{\sqrt{(0.04)(0.96)\left(\frac{1}{250} + \frac{1}{250}\right)}} = 1.38$$

$z_{cal} = 1.38$

<u>level of Significance</u>:- $\alpha = 5\%$

$z_{tab} = 1.96$

<u>Conclusion</u>: $|z_{cal}| < z_{tab}$ $(|1.38| < 1.96)$

Null Hypothesis accepted

There is no diff b/w Sample proportions

UNIT – IV

parameters :- The population measurements of observations are called parameters. It is denoted by θ.

Ex:- μ, σ^2, N

Statistics :- The statistical measurements of observations in a sample are called statistics. It is denoted by $\hat{\theta}$.

Ex:- $\bar{x}$, s^2, n

Estimate:- It is a statement made to find unknown population parameter using sample statistics.

Estimator:- The procedure or rule to determine unknown population parameter using sample statistics. is called an Estimator

Ex:- i) $E(\bar{x}) = \mu$

2) $E(s^2) = \sigma^2$

$\therefore$ $\bar{x}$, s^2 are Estimators.

Types of Estimations :-

1) point Estimation 2) Interval Estimation

1) point estimation:-

If an estimate of the population parameter is given by a single value then that estimate is called point estimation of parameter.

Ex:- The mean height of a student in a

college is 165 cm

Unbiased Estimators :-

A sample statistic $\hat{\theta}$ is said to be unbiased estimator of population parameter θ. If $E[\hat{\theta}] = \theta$.

Ex :- 1) $E(\bar{x}) = \mu$

2) $E(s^2) = \sigma^2$

Biased estimator :-

A sample statistic $\hat{\theta}$ is said to be biased estimator of population parameter θ if $E(\hat{\theta}) \neq \theta$.

Theorem :-

Show that sample mean $\bar{x}$ is unbiased estimator of population mean μ.

proof :-

Let $x_1, x_2, x_3, \ldots x_n$ be a random sample from a population with mean μ.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$E[\bar{x}] = E\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right]$$

we know that $E[x_i] = \mu, \forall i$

$$E[\bar{x}] = \frac{1}{n}E[x_1 + x_2 + x_3 + \ldots x_n]$$

$$= \frac{1}{n}[E(x_1) + E(x_2) + \ldots + E(x_n)]$$

$$= \frac{1}{n}[\mu + \mu + \ldots + \mu]$$

$$= \frac{1}{\cancel{n}}[\cancel{n}\mu]$$

$$\therefore E[\bar{x}] = \mu.$$

Scanned by CamScanner

Theorem:

Show that sample variance $s^2 = \frac{1}{n-1} \sum_{i} (x_i - \bar{x})^2$ is unbiased estimator of population mean σ^2

i.e., $E(s^2) = \sigma^2$.

Proof

$$S^2 = \frac{1}{n-1} \sum' (x_i - \bar{x})^2$$

consider

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} [(x_i - \mu) - (\bar{x} - \mu)]^2$$

$$= \sum_{i=1}^{n} (x_i - \mu)^2 - 2(\bar{x} - \mu) \sum_{i=1}^{n} (x_i - \mu) + \sum_{i=1}^{n} (\bar{x} - \mu)^2$$

$$\Rightarrow \sum_{i=1}^{n} (x_i - \mu)^2 - 2(\bar{x} - \mu) \left(\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \mu \right) + n (\bar{x} - \mu)^2.$$

$$\sum' (x_i - \bar{x})^2 = \sum' (x_i - \mu)^2 - 2(\bar{x} - \mu)(n\bar{x} - n\mu) + n(\bar{x} - \mu)^2$$

$$\Rightarrow \sum' (x_i - \mu)^2 - 2n(\bar{x} - \mu)^2 + n(\bar{x} - \mu)^2 \rightarrow ①.$$

$$S^2 = \frac{1}{n-1} \sum' (x_i - \bar{x})^2$$

$$E(s^2) = \frac{1}{n-1} E \left[\sum' (x_i - \bar{x})^2 \right]$$

$$E(s^2) = \frac{1}{n-1} E \left[\sum_{i=1}^{n} (x_i - \mu)^2 - n(\bar{x} - \mu)^2 \right].$$

$$= \frac{1}{n-1} \left(\sum_{i=1}^{n} E[(x_i - \mu)^2] - n\, E[\bar{x} - \mu]^2 \right).$$

$$= \frac{1}{n-1} \left[\sum_{i=1}^{n} \sigma^2 - n \cdot \sigma_{\bar{x}}^2 \right]$$

$$\Rightarrow \frac{1}{n-1} \left[n\sigma^2 - \cancel{n} \cdot \frac{\sigma^2}{\cancel{n}} \right] \qquad \left[\because \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \right]$$

$$= \frac{1}{n-1} (n-1) \cdot \sigma^2$$

$$\therefore E(s^2) = \sigma^2$$

Note:-

If $s^2 = \frac{1}{n} \sum_{i} (x_i - \bar{x})^2$ then s^2 is biased estimator of σ^2 i.e., $E(s^2) \neq \sigma^2$.

2) The population parameter θ can consists many estimators.

Most efficient parameter Estimator :-

If $\hat{\theta}_1$, $\hat{\theta}_2$ are two unbiased estimators of some population parameter 'θ'. If $V(\hat{\theta}_1) < V(\hat{\theta}_2)$ then $\hat{\theta}_1$ is most efficient estimator of 'θ' than $\hat{\theta}_2$.

(or)

All possible unbiased estimators of parameter θ, the one with smallest variance is called most efficient parameter estimator of θ.

Good Estimator :-

A Good Estimator is the one which is as close to the true value of parameter as possible.

(or)

It satisfies following properties :-

i) consistency :-

Estimator $\hat{\theta}_n$ converges to θ, if $n \to \infty$ then $\hat{\theta}_n$ is consistent.

ii) Unbiasedness :-

The estimator $\hat{\theta}$ is called unbiased if $E[\hat{\theta}] = \theta$.

iii) Efficiency :-

The all possible unbiased estimators of θ, the one with smallest variance is called most efficient estimation.

iv) Sufficiency :-

An estimator is said to be sufficient for a parameter if it contains all the information in the sample regarding parameter.

Interval estimation:-

If an estimate of population parameter is given by a different values. In which the parameter lie. then the estimate is called interval estimation.

Ex:- The mean height of college student is 165 ± 2 cm. ie, (163, 167).

The interval of population parameter θ is of the form $\hat{\theta}$ is $\hat{\theta}_L < \theta < \hat{\theta}_u$ and $P(\hat{\theta}_L < \theta < \hat{\theta}_u) = (1-\alpha)$.

Then $(1-\alpha)$ is called confidence coefficient.

and $(1-\alpha)100\%$ is called confidence percentage.

The interval $\hat{\theta}_L < \theta < \hat{\theta}_u$ is called confidence interval with $(1-\alpha)100\%$ confidence.

Maximum Error & Confidence Interval of μ (Large Sample):-

~~The difference~~

The modulus of difference b/w sample mean and population mean is called maximum error. It is denoted by 'ϵ'.

i.e., $|\bar{x}-\mu| = \text{max. Error} = \epsilon$

$$\epsilon = |\bar{x}-\mu| = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}.$$

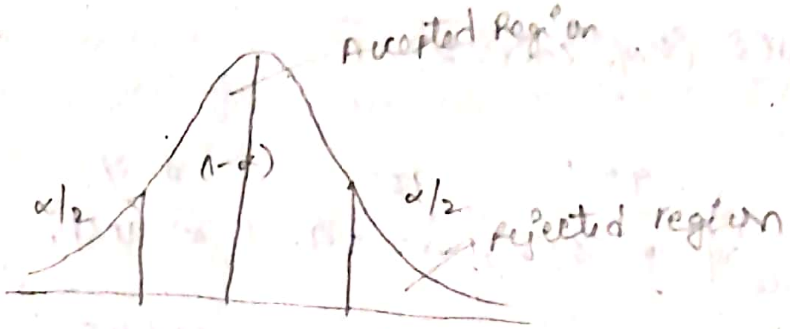
$$\bar{x}-\mu = \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\mu = \bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \longrightarrow (*)$$

$$\bar{x} - \epsilon < \mu < \bar{x} + \epsilon.$$

$$P\left(\bar{x} - z_{\alpha/2} < \mu < \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$



The interval * is called confidence interval for population mean μ with $(1-\alpha)100\%$ confidence.

The confidence $(1-\alpha)$ is given by 99% (or) 95% (or) 90%.

Sample size (n):-

$$e = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\sqrt{n} = \frac{z_{\alpha/2} \cdot \sigma}{e}$$

$$\boxed{n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2}$$

Maximum Error & confidence Interval of μ (small sample):-

(σ is unknown.)

Maximum error is

$$E = |\bar{x} - \mu| = t_{\alpha/2} \cdot \frac{S}{\sqrt{n}}$$

$$\bar{x} - \mu = \pm t_{\alpha/2} \cdot \frac{S}{\sqrt{n}}$$

$$\mu = \bar{x} \pm t_{\alpha/2} \cdot \frac{S}{\sqrt{n}}$$

$$\bar{x} - t_{\alpha/2} \cdot \frac{S}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2} \frac{S}{\sqrt{n}} \rightarrow *$$

$$\bar{x} - E < \mu < \bar{x} + E.$$

Sample size (n):-

$$E = t_{\alpha/2} \cdot \frac{S}{\sqrt{n}}$$

$$\sqrt{n} = \frac{t_{\alpha/2} \cdot S}{E}$$

$$\therefore n = \left(\frac{t_{\alpha/2} \cdot S}{E}\right)^2$$

Table Values ($Z_{\alpha/2}$ (or) Z_α)

$(1-\alpha)$	99%	95%	90%	
level of significance (α) →	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 2\%$
Two-tailed Test	$Z_{\alpha/2} = 2.58$	1.96	1.645	2.33
Right tailed Test	$Z_\alpha = 2.33$	$Z_\alpha = 1.645$	1.28	
Left tailed Test	$Z_\alpha = -2.33$	$Z_\alpha = -1.645$	-1.28	

$Z_{\alpha/2}$ is z-value leaving Area $\alpha/2$ to the right of Normal curve.

1) Find 95% confidence interval for mean of Normal distribution with variance 0.25 using a sample of n=100 values with mean 212.3.

Given $n = 100$

Mean of sample $(\bar{x}) = 212.3$

$S.D = \sigma = \sqrt{0.25} = 0.5$

$(1-\alpha) = 95\%$, $Z_{\alpha/2} = 1.96$

$$\text{Max. Error } (e) = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$= (1.96) \frac{0.5}{\sqrt{100}}$$

$$= 0.098$$

Confidence Interval for μ is

$$\bar{x} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\bar{x} - e < \mu < \bar{x} + e$$

$$212.3 - 0.098 < \mu < 212.3 + 0.098$$

$$212.202 < \mu < 212.398$$

2) A Random sample of size 100 has a std. deviation of 5. what will you say about max. error with 95% confidence.

Given sample size $(n) = 100$

$S.D = \sigma = 5$

$(1-\alpha) = 95\%$, $z_{\alpha/2} = 1.96$

Max. Error $(E) = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$= (1.96)\frac{5}{\sqrt{100}} \Rightarrow 1.96\left(\frac{1}{2}\right) \Rightarrow 0.98$

3) A Random sample of size 100 is taken from a population with $\sigma = 5.1$, Given that sample mean $\bar{x} = 21.6$. Construct 95% confidence limits for population mean ~~the~~ μ.

Given sample size $= 100$.

Mean $(\bar{x}) = 21.6$.

$\sigma = 5.1$

$(1-\alpha) = 95\%$, $z_{\alpha/2} = 1.96$.

Max. Error$(E) = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$= (1.96)\frac{5.1}{\sqrt{100}}$.

$= 0.9996$

Confidence Interval for μ is

$\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$.

$\bar{x} - E < \mu < \bar{x} + E$.

$21.6 - \overset{0.996}{<} \mu < 21.6 - 0.9996$

$\Rightarrow$

4) Assume that $\sigma = 20$, How large a Random Sample be taken to assert with probability 0.95 that the sample mean will not differ from the true mean by more than 3 points.

Given Max-Error $(E) = 3$.

$\sigma = 20$

$(1-\alpha) = 0.95 = 95\%$

$z_{\alpha/2} = 1.96$

$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$.

$n = \left(\frac{z_{\alpha/2}\,\sigma}{E}\right)^2$

$= \left[\frac{(1.96)20}{3}\right]^2 = (13.06)^2 = 26.12$

5) In a Study of Automobile Insurance a random sample of 80 body repair cost had mean of ₹ 472.36 and S.D of ₹ 62.35. If $\bar{x}$ is used as a point estimate to the true average repair cost with what confidence, we can assert that the max-error does not exceed ₹ 10/-

Max-Error $(E) = 10/-$

Sample size $(n) = 80$

Sample Mean $(\bar{x}) = 472.36$,

$\sigma = 62.35$

$z_{\alpha/2} = ?$

$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$.

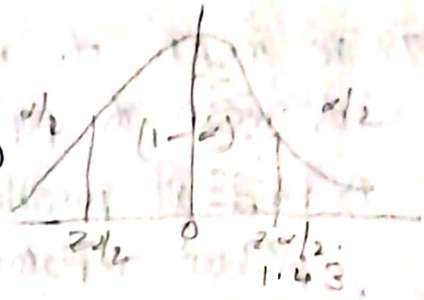
$10 = z_{\alpha/2} \cdot \frac{62.35}{\sqrt{80}}$,

$z_{\alpha/2} = \frac{10 \times \sqrt{80}}{62.35}$

$= 1.43$

from Tables,

$P(0 < Z < 1.43) = \frac{1-\alpha}{2} = 0.4236.$



confidence $= (1-\alpha) = 2 \times (0.4236)$

$= 0.8472.$

$\Rightarrow 84.72\%.$

6) The efficiency expert of computer company tested 40 engineers to estimate the average time it takes to assemble a certain computer component, getting mean of 12.73 mins and S.D 2.06 mins

i) If $\bar{x}$ is 12.73 mins is used as point estimate of Actual average time required to perform the task, find max. Error with 99% confidence.

ii) Construct 98% confidence interval for true average time it takes to do the job.

iii) With what confidence can be assert that the sample mean doesnot differ from the true mean by more than 30 secs.

$e = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}.$

$0.5 = Z_{\alpha/2} \cdot \frac{2.06}{\sqrt{40}}$

$Z_{\alpha/2} = \frac{0.5(\sqrt{40})}{2.06} = 1.535.$

from tables.

$P(0 < Z < 1.535) = \frac{1-\alpha}{2} = 0.4382$

confidence $= (1-\alpha) = 2 \times (0.4382)$

$= 0.8764.$

$= 87.64\%.$

7) The mean of random sample is unbiased estimate of mean of population 3, 6, 9, 15, 27.

i) Find all possible samples of size '3'. that can be taken without replacement from finite population.

ii) Calculate the mean. of each of samples in (i) and assigning each sample a probability of 1/10. Verify that mean of these $\bar{x}$ is equal to 12. and verify $E(\bar{x}) = \mu$.

i) Population $= \{3, 6, 9, 15, 27\}$

population size $(N) = 5$

Sample size $(n) = 3$.

No. of samples without replacement $= {}^{N}C_{n} = {}^{5}C_{3}$

$= \frac{5 \times 4 \times 3}{3 \times 2 \times 1} = 10.$

Samples $= \{(3,6,9), (3,6,15), (3,6,27), (3,9,15), (3,9,27), (3,15,27), (6,9,15), (6,9,27), (9,15,27), (6,15,27)\}$.

Sample Mean $= \{6, 8, 12, 9, 18, 15, 10, 14, 17, 16\}$.

Population Mean $(\mu) = \frac{3+6+9+15+27}{5} = \frac{60}{5}$.

$\mu = 12$.

Assigning prob. 1/10 to each sample mean.

$\bar{x}$	6	8	9	10	12	13	14	15	16	17.
$P(\bar{x})$	$\frac{1}{10}$	$\frac{1}{10}$	.	.	.	.	.	.	.	.

$E(\bar{x}) = \sum x_i P(x_i)$

$= 6\left(\frac{1}{10}\right) + 8\left(\frac{1}{10}\right) + \cdots + 17\left(\frac{1}{10}\right)$

$\Rightarrow \frac{1}{10}[6 + 8 + \cdots + 17]$

$\Rightarrow \frac{120}{10}$

$\Rightarrow 12$

8) Suppose that we observe a R.V having binomial distribution and getting x success in n trails.

i) show that x/n is an unbiased estimate of binomial parameter 'p'.

ii) Show that $\frac{x+1}{n+2}$ is not unbiased estimate of binomial parameter 'p'.

9) Find 95% confidence limit for the mean of normal distributed pop. from which the sample was taken 15, 17, 10, 18, 16, 9, 7, 11, 13, 14.

8) i) If X is R.V having Binomial dist. $mean(\mu) = E(x)$
$= np$

$$E\left[\frac{x}{n}\right] = \frac{E(x)}{n} = \frac{\cancel{n}p}{\cancel{n}} = p$$

ii) $$E\left[\frac{x+1}{n+2}\right] = \frac{1}{n+2} E[x+1] \Rightarrow \frac{1}{n+2}[E(x)+1].$$

$$\Rightarrow \frac{np+1}{n+2} \neq p.$$

$\therefore \frac{x+1}{n+2}$ is ~~so~~not unbiased estimator of p.

9) Confidence limits

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} \quad [\sigma \text{ is unknown}]$$

Sample ~~mean~~ variance

$$s^2 = \frac{1}{n-1} \Sigma(x_i - \bar{x})^2$$

Sample = {15, 17, 10, 18, 16, 9, 7, 11, 13, 14}

$$\text{Sample mean } (\bar{x}) = \frac{15+17+10+18+16+9+7+11+13+14}{10}$$

$$= \frac{130}{10}$$

$$\Rightarrow 13$$

$$S^2 = \frac{1}{10-1}\{(15-13)^2+(17-13)^2+\cdots\cdots+(14-13)^2\}$$

$$= \frac{1}{9}[4+16+9+25+9+16+36+4+0+1].$$

$$= \frac{120}{9} \Rightarrow \frac{40}{3}.$$

$$\boxed{S = 3.65}$$

$(1-\alpha)=95\%$, $z_{\alpha/2}=1.96$, $n=10$.

confidence interval is $13 \pm (1.96)\frac{(3.65)}{\sqrt{10}}$

$\Rightarrow 13 \pm 2.262$

$\Rightarrow (13-2.262,\ 13+2.262)$

$\Rightarrow (10.74,\ 15.26)$.

Hypothesis :-

The value of parameter whether to accept or reject a statement about parameter then that statement is called hypothesis.

Ex:- i) The majority of men in the city are smokers

ii) A drunk chemist is to decide whether a new drug is really effective in curing a desease.

Test of Hypothesis:-

The procedure which enables us to decide on the basis of sample results whether a hypothesis is true or not, is called test of Hypothesis.

2m

** Null Hypothesis :-

A hypothesis of no difference is called null hypothesis.

It is denoted by H_0 and defined as

$$H_0 : \mu = \mu_0$$

Alternative Hypothesis :-

The opposite statement of null hypothesis is called alternative hypothesis. It is denoted by H1. defined as

i) $H_1 : \mu \neq \mu_0$ ($\mu > \mu_0$ (or) $\mu < \mu_0$) (Two-tailed).

ii) $H_1 : \mu > \mu_0$ (Right tailed)

iii) $H_1 : \mu < \mu_0$ (Left tailed).

Level of significance(α) :-

It is the confidence with which we reject or accept Null hypothesis Ho. 5% of level of significance in a test procedure indicates that there are 5 cases in 100 that reject Null hypothesis.

*** Types of Errors :-

1) Type-I Error :- (Reject Ho when it is True).

If the Null hypothesis is Ho is true but it rejected by test procedure. Then the error made is called Type-I Error (or) α error.

P(Reject Ho when it is True) = α - Error.

It is also called procedure's risk.

2) Type-II Error :- (Accept Ho when it is false).

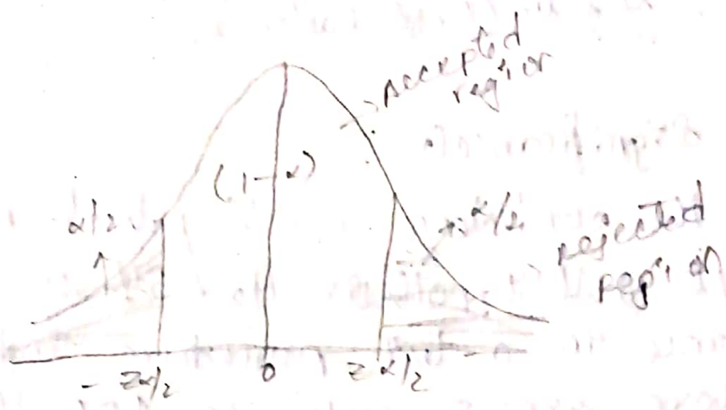
If the Null hypothesis is Ho is false but it is accepted by test procedure, Then the error is called Type-II error (or) β error.

It is also called consumer's risk.

Critical Region :-

A region corresponding to statistic t in the sample space S which leads to the rejection of H_0 is called critical region (or) Rejection region.

The region which leads to acceptance of H_0 is called acceptance region.



Critical values :-

The value of test statistic which seperates the critical region and acceptance region is called critical value (or) significant value.

** Two tailed Test & one-tailed test.

i) Two tailed test :-
If the alternative hypothesis of the form $H_1 : \mu \neq \mu_0$ ($\mu > \mu_0$ (or) $\mu < \mu_0$) in a test of hypothesis then the test process is called two-tailed test.

ii) Right tailed test :-

If the alternative hypothesis of the form $H_1 : \mu > \mu_0$ in a test of hypothesis then the test is called Right Tailed test.

iii) Left Tailed Test :-

If the alternative hypothesis of the form $H_1 : \mu < \mu_0$ in a test of Hypothesis then it is called Left tailed test

Process of Testing of Hypothesis :-

step 1> Null Hypothesis :- $H_0 : \mu = \mu_0$

A hypothesis with no difference is called Null hypothesis.

step 2> Alternative hypothesis :-

It is defined as one of the following

i> $H_1 : \mu \neq \mu_0$ ($\mu > \mu_0$ (or) $\mu < \mu_0$)

ii> $H_1 : \mu > \mu_0$

iii> $H_1 : \mu < \mu_0$.

step-3> Level of Significance :-

usually it is 1% (or) 5% (or) 10%.

Table value = $z_{\alpha/2}$ (or) z_α

step 4> Test statistic :-

$$z = \frac{t - E(t)}{S.E(t)}$$

where t - statistic

step 5> Conclusion :-

If $|z| < z_{\alpha/2}$ then H_0 is true (or) H_0 is accepted.

If $|z| > z_{\alpha/2}$ then H_0 is false (or) H_0 is rejected. so, the H_1 is true.

critical values :-

	(α) level of significance			
	1%	5%	10%	2%
Two-tailed	$\lvert z_{\alpha/2}\rvert = 2.58$	1.96	1.645	2.33
Right-tailed	$z_\alpha = 2.33$	1.645	1.28	
Left tailed	$z_\alpha = -2.33$	-1.645	-1.28	

Large sample :-

Test of significance of sample mean :-

A Random sample of size n has the sample mean $\bar{x}$, which is taken from the population with mean μ and S.D σ and the population mean μ has specified value μ_0

i) Null Hypothesis :- $H_0: \mu \neq \mu_0$

2) Alternative Hypothesis :-

i) $H_1 : \mu \neq \mu_0$

ii) $H_1 : \mu > \mu_0$

iii) $H_1 : \mu < \mu_0$.

3) Level of significance :-

Table value $= z_{\alpha/2}$.

4) Test statistic :-

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

5) Conclusion :-

if $|z| < z_{\alpha/2}$ H_0 True

if $|z| > z_{\alpha/2}$ H_0 false

1) Sample of 400 items are taken from a population whose std. deviation 10. The mean of sample is 40. Test whether the sample has come from a population with mean 38. Also calculate 95% confidence limits.

Sample size $(n) = 400$
Sample mean $(\bar{x}) = 40$
S.D population $(\sigma) = 10$
Population mean $(\mu) = 38$

i) Null Hypotheses :- $\mu = 38$

ii) Alternative Hypothesis :- $H_0 : \mu \neq 38$

iii) Level of significance :-

$(1-\alpha) = 95\% \Rightarrow \alpha = 5\%$

Table value $= z_{\alpha/2} = 1.96$ (Two-tailed)

iv) Test statistic :-

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \Rightarrow \frac{40 - 38}{10/\sqrt{400}} \Rightarrow \frac{2 \times 20}{10} = 4.$$

v) Conclusion :-

$z = 4$, $z_{\alpha/2} = 1.96$

$|z| > z_{\alpha/2}$

$\therefore$ H_0 is false

$\therefore$ H_1 is true

$\mu \neq 88$ //

confidence limits with 95% for μ is

$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ (or) $\bar{x} \pm \epsilon$.

$\epsilon = (1.96)\frac{10}{\sqrt{400}} \Rightarrow (1.96)\frac{10}{20}$

$\Rightarrow \frac{1.96}{2} \Rightarrow 0.98$

confidence limits.

$(\bar{x} - \epsilon., \bar{x} + \epsilon)$

$(40 - 0.98, 40 + 0.98)$

$(39.02 + 40.98)$

2) A Sample of 900 items has mean 3.4 and S.D 2.61. If this sample has been taken from a large population of mean 3.25 and S.D 2.61. If the population is normal and find 95% confidence limits for mean.

3) An Ambulance service claims that it takes on the average less than 10 mins to reach its destination in emergency calls. A sample of 36 calls has mean of 11 mins. and variance 16 mins. Test the claim and 5% level of significance

4) A sample of 64 students have mean weight 70 kgs. can this be regarded as a sample from population with mean weight 56 kgs and std-deviation 25 kgs.

2) Given

Sample size(n) = 900

Sample mean ($\bar{x}$) = 3.4

population s.D(σ) = 2.61

population mean(μ) = 3.25.

i) Null Hypothesis :-

$H_0 : \mu = 3.25$

ii) Alternative Hypothesis.

$H_1 : \mu \neq 3.25$

Two tailed.

iii) Level of significance :-

$(1-\alpha) = 95\% \Rightarrow \alpha = 5\%$

Table value $Z_{\alpha/2} = 1.96$

iv) Test statistic :-

$$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{3.4-3.25}{2.61/\sqrt{900}} = 1.724.$$

v) Conclusion :-

$z = 1.724$, $z_{\alpha/2} = 1.96$

$|z| < z_{\alpha/2}$

H_0 is True

$\therefore \mu = 3.25$.

Confidence limits :-

$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

(or)

$\bar{x} \pm e$

Max. Error (e) $= z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$= (1.96) \frac{2.61}{\sqrt{900}}$

$e = 0.17$

confidence interval is $= (\bar{x}-e, \bar{x}+e)$

$= (3.4-0.17, 3.4+0.17)$

$= (3.23, 3.57)$.

3) Given,

Sample size $(n) = 36$ calls.

Sample's mean $(\bar{x}) = 11$ mins.

Population S.D $(\sigma^2) = 16$ mins $\Rightarrow \sigma = 4$ mins.

Population mean $(\mu) = 10$ mins.

i) Null Hypothesis :-

$H_0 : \mu = 10$ mins.

ii) Alternative Hypothesis :-

$H_1 : \mu < 10$ mins.

(left tailed)

iii) Level of significance :-

$(1-\alpha) = 95\% \Rightarrow \alpha = 5\%$.

Table value $z_\alpha = -1.645$.

iv) Test Statistic :-

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \Rightarrow \frac{11-10}{4/\sqrt{36}} = 1.5.$$

v) Conclusion :-

$z = 1.5$, $z_\alpha = -1.645$

$|z| = 1.5$, $|z_\alpha| = 1.645$.

$|z| < |z_\alpha|$

$\therefore H_0$ is True

$\mu = 10$ mins

4) Given

Sample size $(n) = 64$

Sample mean $(\bar{x}) = 70$ kgs

population S.D $(\sigma) = 35$ kgs

population mean $(\mu) = 56$ kgs.

i) Null Hypothesis :-

$H_0 : \mu = 56$.

ii) Alternative Hypothesis :-

$H_1 : \mu \neq 56$

Two tailed.

iii) Level of significance.

$\alpha = 5\%$.

Table value $Z_{\alpha/2} = 1.96$.

iv) Test statistic :-

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \Rightarrow \frac{70-56}{25/\sqrt{64}}.$$

$z = 4.48$

v) Conclusion :-

$Z = 4.48$ $\quad Z_{\alpha/2} = 1.96$

$|Z| > Z_{\alpha/2}$.

$\therefore H_0$ is false

$\therefore H_1$ is True

$\mu \neq 56$ kg's.

Test of Equality of Two means (or) Test of significance difference b/w two means.

i) Null Hypothe

Let $\bar{x}$ and $\bar{y}$ be sample means of two independent large samples sizes n_1 and n_2 drawn from two populations having means μ_1 and μ_2 with the std. deviations σ_1 & σ_2. To Test whether the two population means are equal,

i) Null Hypothesis :-

$H_0 : \mu_1 = \mu_2$ (or) $\mu_1 - \mu_2 = 0$.

ii) Alternative Hypothesis :-

i) $H_1 : \mu_1 \neq \mu_2$ ($\mu_1 > \mu_2$ (or) $\mu_1 < \mu_2$) (Two tailed)

ii) $H_1 : \mu_1 > \mu_2$ (Right tailed)

iii) $H_1 : \mu_1 < \mu_2$ (Left tailed)

iii) Level of significance :- (α)

Table value = $Z_{\alpha/2}$ (or) Z_α

iv) Test statistic :-

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

if $\sigma_1 = \sigma_2 = \sigma$ (say)

$$Z = \frac{\bar{x} - \bar{y}}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}.$$

v) Conclusion :-

if $|Z| < Z_{\alpha/2}$ $\therefore H_0$ True

if $|Z| > Z_{\alpha/2}$ $\therefore H_0$ false

$\therefore H_1$ is True

1) The means of two large samples of size s. $n_1 = 1000$ and 2000 members are 67.5 ~~and~~ inches and 68 inches respectively. Can the samples be regarded as drawn from the same population of S.D 2.5 inches

Given,

Sample mean $\bar{x}$ = 67.5 inches $\bar{y}$ = 68 inches

sample sizes $n_1 = 1000$ $n_2 = 2000$

population S.D = 2.5 inches.

i) Null type :-

$H_0 : \mu_1 = \mu_2$

ii) Alternative Hypothesis :-

$H_1 : \mu_1 \neq \mu_2$ (Two-tailed)

iii) Level of significance (α) = 5%.

Table value t = $Z_{\alpha/2}$ = 1.96

iv) Test Statistic :-

$$z = \frac{\bar{x} - \bar{y}}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{67.5 - 68}{(2.5)\sqrt{\frac{1}{1000} + \frac{1}{2000}}} \Rightarrow -5.16$$

v) Conclusions :-

$Z = -5.16$ $\quad Z_{\alpha/2} = 1.96$

$|z| = 5.16$ $\quad Z_{\alpha/2} = 1.96$

$|z| > Z_{\alpha/2}$

$\therefore H_0$ is false

$\therefore H_1$ is True

$\mu_1 \neq \mu_2$

$\therefore$ Two samples has drawn from 2 different populations

2) A Researcher wants to know the intelligence of students in a school is selected 2 groups of students, in the first group there are 150 students having mean IQ of 75 with S.D 15, in the second group there are 250 students having mean IQ of 70 with S.D 20. Test whether the 2 groups of students have taken from same school.

3) The average marks scored by 32 boys is 72 with S.D 8. while that for 36 girls is 70 with S.D 6. Does this indicate that boys perform better than girls at 5% level of significance.

4) The mean height of 50 male students who participated in sports is 68.2 inches with S.D 2.5. The mean height of 50 male students who have not participated in sports is 67.2 inches with S.D 2.8. Test the hypothesis the height of students who participated in sports is more than the students who have not participated in sports.

2) 1^{st} group => Sample size (n_1) = 150
Sample mean ($\bar{x}$) = 75
population S.D (σ_1) = 15

2^{nd} group => Sample size (n_2) = 250
Sample mean ($\bar{x}$) = 70
population ~~size~~ (σ_2) = 20.

i) Null Hypothesis :-
$H_1 : \mu_1 = \mu_2$.

ii) Alternative Hypothesis.
$H_1 : \mu_1 \neq \mu_2$.

iii) Level of significance
$\alpha = 5\%$.
Table value $z_{\alpha/2} = 1.96$

iv) Test ~~of~~ statistic :-

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \Rightarrow \frac{75 - 70}{\sqrt{\frac{(15)^2}{150} + \frac{(20)^2}{250}}} \Rightarrow 2.83$$

v) Conclusion :-
$z = 2.83$, $z_{\alpha/2} = 1.96$
$|z| > z_{\alpha/2}$
$\therefore H_0$ is False (Rejected)
$\therefore H_1$ is True
$\mu_1 \neq \mu_2$

3) Boys :-
Sample size (n_1) = 32
Sample mean ($\bar{x}$) = 72
population S.D (σ_1) = 8

Girls
Sample size (n_2) = 36
Sample mean ($\bar{y}$) = 70
population S.D (σ_2) = 6.

Scanned by CamScanner

i) Null Hypothesis :-

$H_0 : \mu_1 = \mu_2$

ii) Alternative Hypothesis :-

$H_1 : \mu_1 > \mu_2$.

iii) Level of significance (α)

$\alpha = 5\% = 0.05$

Table value $Z_{\alpha/2} = 1.645$.

iv) Test significance :-

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \Rightarrow \frac{72 - 70}{\sqrt{\frac{(8)^2}{32} + \frac{6^2}{36}}} \Rightarrow 1.154.$$

v) Conclusion :-

$Z = 1.154$ $\quad Z_{\alpha/2} = 1.645$

$|Z| < Z_{\alpha/2}$

$\therefore H_0$ is True (Accepted).

$\mu_1 = \mu_2$.

$\therefore$ Both boys & girls perform equally.

4) Male in sports.

Sample size $(n_1) = 50$

Sample mean $(\bar{x}) = 68.2$.

Population S.D $(\sigma_1) = 2.5$

Male not in sports.

Sample size$(n_2) = 50$

Sample mean$(\bar{x}) = 67.8$.

population S.D$(\sigma_2) = 2.8$.

i) Null Hypothesis :-

$H_0 : \mu_1 = \mu_2$

ii) Alternative Hypothesis :-

$H_1 : \mu_1 > \mu_2$

iii) Level of significance (α)

$\alpha = 5\% = 0.05$.

Table value $Z_{\alpha/2} = 1.645$

iv) Test statistic :-

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \Rightarrow \frac{68.2 - 67.2}{\sqrt{\frac{(2.5)^2}{50} + \frac{(2.8)^2}{50}}} \Rightarrow 1.88.$$

v) Conclusion.

$z = 1.88$ $\quad Z_{\alpha/2} = 1.645$

$|z| > Z_{\alpha/2}$

$\therefore H_0$ is False (Rejected)

$\therefore H_1$ is True.

$\mu_1 > \mu_2$.

Test of significance of single proportion

Suppose a large Random sample of size 'n' has a sample proportion 'p' which is taken from a population proportion 'P'.

$$\text{proportion} = \frac{\text{No. of Success}}{\text{Total observations}} = \frac{x}{n}$$

Sample proportion $= p = \frac{x}{n}$

Population proportion $= P$.

To test the population proportion 'P' which has specified value P_0.

i) Null Hypothesis :-

$H_0 : P = P_0$

ii) Alternative Hypothesis :-

i) $H_1 : P \neq P_0$ ($P > P_0$ (or) $P < P_0$).

ii) $H_1 : P > P_0$

iii) $H_1 : P < P_0$

iii) Level of significance :- (α)

Table value $= Z_{\alpha/2}$ (or) Z_{α}

iv) Test statistic.

$$Z = \frac{p-P}{\sqrt{\frac{PQ}{n}}} . \qquad Q = 1-P.$$

v) Conclusion :-

if $|Z| < Z_{\alpha/2}$. Ho True

if $|Z| > Z_{\alpha/2}$ Ho False.

Confidence Limit for P :-

1) $p - Z_{\alpha/2} \cdot \sqrt{\frac{PQ}{n}} < P < p + Z_{\alpha/2} \sqrt{\frac{PQ}{n}}$

(or)

$$p - 3\sqrt{\frac{PQ}{n}} < P < p + 3\sqrt{\frac{PQ}{n}}$$

if P is unknown.

$$p - 3\sqrt{\frac{pq}{n}} < P < p + 3\sqrt{\frac{pq}{n}}$$

1) In a sample of 1000 people in Karnataka 540 are rice eaters. and the rest are wheat. Can we assume that both rice and wheat eaters are equally popular in this state at ~~one~~ 1% level of significance

Sample size (n) = 1000

No. of rice eaters (x) = 540

Sample proportion (p) $= \frac{x}{n} = \frac{540}{1000} = 0.54$.

population proportion (P) = 50%.

$P = 0.5$.

$Q = 1 - P = 1 - 0.5 = 0.5$

i) Null Hypothesis :-

Ho : $P = 50\% = 0.5$.

2) Alternative Hypothesis.

$H_1 : P \neq 0.5$.

3) Level of significance (α)

$\alpha = 1\%$.

Table value $= Z_{\alpha/2} = 2.58$.

4) Test statistic :-

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} \Rightarrow \frac{0.54 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{1000}}} \Rightarrow 2.532 ,$$

5) Conclusion :-

$Z = 2.532$, $Z_{\alpha/2} = 2.58$

$|Z| > Z_{\alpha/2}$

$\therefore H_0$

2) In a big city ~~out of 600 men~~ 325 men out of 600 men were found to be smokers. Does this information support the conclusion that the majority of men in the city are smokers.

3) A manufacturer claimed that 95% of equipment which he supplied to a factory conformed to specifications. And examination of a sample of 200 pieces of equipment revealed that 18 were faulty. Test its claim at 5% level of significance.

4) Among 900 people in a state 90 are found to be chapathi eaters. Construct 99% confidence interval for true population.

5) The experience had shown that 20% of a manufactured product is of top quality. In one days production of 400 articles only 50 are top quality. Test the Hypothesis at 0.05 level.

2) Sample size of men $(n) = 600$

No. of men Smokers $(x) = 325$.

sample proportion of smokers $(p) = \frac{x}{n} = \frac{325}{600}$

$p = 0.5417$

population proportion $(P) = 50\%$.

i) Null Hypothesis :- $P = 50\%$.

ii) Alternative Hypothesis :- H_1 ; $P > 50\%$.

iii) Level of Significance :- (α)

$\alpha = 5\%$

Table value $= Z_\alpha = 1.645$.

iv) Test statistic.

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} \Rightarrow \frac{0.5417 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{600}}} \Rightarrow 2.04.$$

v) conclusion :-

$Z = 2.04$, $Z_\alpha = 1.645$.

$|Z| > Z_\alpha$.

$\therefore$ Ho is false

$\therefore$ H_1 is True.

$P > 50\%$.

majority of men are smokers.

3) Sample size $(n) = 200$

No. of faulty $= 18$

No. of good condition items $(x) = 200 - 18$

$= 182$

Sample proportion of good $(p) = \frac{x}{n} = \frac{182}{200} = 0.91$

population proportion$(P) = 95\% = 0.95$

$Q = 1 - P$.

$\Rightarrow 1 - 0.95$

$\Rightarrow 0.05$.

i) Null Hypothesis :- $p = 95\%$.

ii) Alternative Hypo :-

$H_1 : p \neq 95\%$.

iii) Level of significance (α)

$\alpha = 5\%$.

Table value $= z_{\alpha/2} = 1.96$.

iv) Test statistic :-

$$z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} \Rightarrow \frac{0.91 - 0.95}{\sqrt{\frac{(0.95)(0.05)}{200}}} \Rightarrow -2.59$$

v) Conclusion :-

$z = -2.59$, $z_{\alpha/2} = 1.96$

$|z| = 2.59$, $z_{\alpha/2} = 1.96$.

$|z| > z_{\alpha/2}$.

$\therefore H_0$ is False

H_1 is True

$p \neq 95\%$.

4) Sample size (n) $= 900$

No. of chapathi eaters (x) $= 90$.

sample proportion of chapathi eaters $= p = \frac{x}{n} = \frac{90}{900}$

$= 0.1$

$q = 1 - p = 1 - 0.1 = 0.9$

$(1-\alpha) = 99\%$, $z_{\alpha/2} = 2.58$.

Confidence Interval is

$$p - z_{\alpha/2} \sqrt{\frac{pq}{n}} < P < p + z_{\alpha/2} \sqrt{\frac{pq}{n}}.$$

$$0.1 - 2.58 \sqrt{\frac{0.1 \times 0.9}{900}} < P < 0.1 + 2.58 \sqrt{\frac{0.1 \times 0.9}{900}}$$

$\Rightarrow 0.07 < P < 0.13$.

$\therefore (0.07, 0.13)$

5) Sample size (n) = 400

no. of top quality (x) = 50.

Sample proportion of top quality (p) = $\frac{x}{n} \Rightarrow \frac{5\cancel{0}}{40\cancel{0}} \Rightarrow 0.125$

population proportion of top quality (P) = 20% = 0.2.

$Q = 1 - 0.2 \Rightarrow 0.8$

i) Null Hypothesis :-

$P = 20\%$

ii) Alternative Hypothesis

$H_1 : P \neq 20\%$

iii) Level of significance :- (α)

$\alpha = 5\%$

Table value $= z_{\alpha/2} = 1.96$.

iv) Test statistic :-

$$z = \frac{p-P}{\sqrt{\frac{PQ}{n}}} \Rightarrow \frac{0.125.}{\sqrt{\frac{(0.2)(0.8)}{400}}} \Rightarrow -3.75.$$

v) Conclusion :-

$z = -3.75, \; z_{\alpha/2} = 1.96$

$|z| = 3.75, \; z_{\alpha/2} = 1.96$

$|z| > z_{\alpha/2}$

H_0 is false

H_1 is true.

$P \neq 20\%$

note :

Test statistic

$$z = \frac{p-P}{\sqrt{\frac{PQ}{n}}} \quad \text{where } \sqrt{\frac{PQ}{n}} \text{ is standard Error.}$$

Max Error $= z_{\alpha/2}\sqrt{\frac{PQ}{n}}$

sample proportion

$$z. = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \quad \text{where } \frac{\sigma}{\sqrt{n}} \text{ is S.E.}$$

Max-Error $= z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$.

Test of significance of difference b/w two proportions.

let p_1, p_2 be two sample proportions of size's n_1, n_2 drawn from population proportions P_1, P_2. To test whether the two samples have been drawn from sample population,

$$p_1 = \frac{x_1}{n_1}$$

$$p_2 = \frac{x_2}{n_2}.$$

i) Null Hypothesis:

$$H_0 : P_1 = P_2$$

ii) Alternative Hypothesis:-

i) $H_1 : P_1 \neq P_2$

ii) $H_1 : P_1 > P_2$

iii) $H_1 : P_1 < P_2$.

iii) Level of significance

iv) Test Statistic.

case (i) if $P_1 - P_2 = 0$

$$z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2}$.

$$Q = 1 - P.$$

case (ii) if $P_1 - P_2 \neq 0$

$$z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$Q = 1 - P.$$

v) Conclusion :-

If $|z| < Z_{\alpha/2}$ H_0 True

$|Z| > Z_{\alpha/2}$ H_1 True

Standard Error $(P_1 - P_2)$ is

$$SE(P_1 - P_2) = \sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Test UNIT-IV Test Statistic

Large sample tests ($n \geq 30$)

1) Test of significance of single mean

$\bar{x}, n$

Table value $= Z_{\alpha/2}$ (or) Z_α

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$H_0 : \mu = \mu_0$
$H_1 : \mu \neq \mu_0$

2) Test of sign. difference b/w Two means

$\bar{x}$ $\bar{y}$
n_1 n_2
σ_1 σ_2

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

3) Test of sign. of single proportion

$\bar{x}, \sigma \rightarrow$ not given.

$p = \frac{x}{n}$

$$z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$$

$Q = 1 - P$

$H_0 : P = P_0$
$H_1 : P \neq P_0$

4) Test of sign. difference b/w two proportions.

$\bar{x}, \sigma \rightarrow$ not given.

x_1 x_2
n_1 n_2

$p_1 = \frac{x_1}{n_1}$ $p_2 = \frac{x_2}{n_2}$

$$z = \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$H_0 : P_1 = P_2$
$H_1 : P_1 \neq P_2$.

where p = combined proportion.

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2}$$

$q = 1 - p$.

Small Sample Test ($n < 30$)

1) t-test

Test of significance of single mean.

$\bar{x}, n, v = (n-1)$

case i) if 's' is given

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}}$$

case ii) is s is calculated (S) i.e, s is not given directly

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n}}$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

for Two tailed:

Table value $= t_{\alpha/2}$

for one-tailed:

Table value $= t_\alpha$.

2) Test of sign. difference b/w two means.

$\bar{x}$ $\bar{y}$

n_1 n_2

s_1^2 s_2^2

for t

degrees of freedom:

$$v = n_1+n_2-2.$$

Two tailed test:-

Table value $= t_{\alpha/2}$

one tailed test:-

Table value $= t_\alpha$

3) paired t-test:-

x_i	Before	
y_i	After.	

$d_i = x_i - y_i$ (or) $d_i = y_i - x_i$

$d.f\ (v) = (n-1)$

F- Tes

case i) if s_1, s_2 are given directly.

$$S^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}.$$

$$t = \frac{\bar{x}-\bar{y}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}.$$

case ii) if $s_1\ s_2$ are not given directly.

$$S^2 = \frac{1}{n_1+n_2-2}\left[\sum_{i=1}(x_i-\bar{x})^2 + \sum_{i}(y_j-\bar{y})^2\right]$$

$\bar{x} = \frac{1}{n_1}\sum x_i$, $\bar{y} = \frac{1}{n_2}\sum y_j$.

$$t = \frac{\bar{x}-\bar{y}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}.$$

$$t = \frac{\bar{d}}{S/\sqrt{n}}$$

$$\bar{d} = \frac{1}{n}\sum_{i} d_i$$

$$S^2 = \frac{1}{n-1}\sum_{i}(d_i-\bar{d})^2$$

F - test :-

Variances

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

$$F = \frac{\text{Greater variance}}{\text{smaller variance}}$$

$$F = \frac{S_1^2}{S_2^2}, \text{ if } S_1^2 > S_2^2.$$

$$S_1^2 = \frac{n_1 s_1^2}{n_1 - 1} = \frac{\sum (x_i - \bar{x})^2}{n_1 - 1}$$

$$S_2^2 = \frac{n_2 s_2^2}{n_2 - 1} = \frac{\sum (y_i - \bar{y})^2}{n_2 - 1}$$

Table value :-

$F_\alpha (\nu_1, \nu_2)$

$\nu_1 = (n_1 - 1)$

$\nu_2 = (n_2 - 1)$

χ^2 - test

1) Goodness of fit :-

$H_0 : O_i = E_i$

$H_1 : O_i \neq E_i$

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Table value :-

$d.f (\nu) = n - 1$

χ^2_α

2) Independence of attributes.

H_0 : there are no association b/w attributes.

H_1 : There is a association b/w attributes.

	C	D
A	a	b
B	c	d

(O_i)

$\Rightarrow$

	C	D	
A	$\frac{(a+b)(a+c)}{N}$	$\frac{(b+d)(a+b)}{N}$	$a+b$
B	$\frac{(a+c)(c+d)}{N}$	$\frac{(b+d)(c+d)}{N}$	$c+d$
	$a+c$	$b+d$	$N = a+b+c+d$

(E_i)

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

$d.f (\nu) = (\text{No. of rows} - 1) \times (\text{No. of columns} - 1)$

Table value $= \chi^2_\alpha$

Scanned by CamScanner

UNIT-V

Stochastic processes and Markov Chains

Short Answer questions

(1) Define Stochastic process, Markov Chain & Transition Matrix

A:- Stochastic process:- The family of all the random variables at particular time 't' is known as Stochastic process.

Ex:- A Queuing system, turblent fluid flow.

Markov chain:- A Stochastic process is said to be Markov process (or) chain if it satifies markov property. i.e if occurance future state is depends on present state.

i.e $P\{X_{n+1} = x_{n+1} \mid X_n = x_n\}$.

Transition matrix:- The probability of future state is depends on present state is known as Transition matrix.

i.e $P\{X_{n+1} = j \mid X_n = i\} = P_{ij}$

(2) If the transition probability matrix is $\begin{bmatrix} 0 & 0.2 & x \\ x & 0.1 & y \\ 0.1 & 0.2 & z \end{bmatrix}$ find x, y, z

A:- The matrix is said to be Transition probability

Matrix if it satifies following conditions

(a) It is a square matrix with non-negative elements.

(b) Sum of each row is equal to '1'

$\therefore \rightarrow 0+0.2+x=1 \Rightarrow x=0.8$

$\rightarrow x+0.1+y=1$

$0.8+0.1+y=1$

$y=0.1$

$\rightarrow 0.1+0.2+z=1$

$z=0.7$

$\therefore x=0.8,\ y=0.1,\ z=0.7$

(3) Define Regular Stochastic process matrix with Example.

Ans: A Matrix is said to be regular stochastic if some powers of 'p' becomes non-zero elements of matrix.

Ex:- $A=\begin{bmatrix} 0 & 1 \\ 1/4 & 3/4 \end{bmatrix}$, $A^2=\begin{bmatrix} \frac{1}{4} & 3/4 \\ 3/16 & 13/16 \end{bmatrix}$

$\therefore$ A is regular Matrix

(4) Find the equilibrium vector of $\begin{bmatrix} 1/4 & 3/4 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

Sol:- Given that $P = \begin{bmatrix} 1/4 & 3/4 \\ 1/2 & 1/2 \end{bmatrix}$

Let X be the probability vector. We want to find X such that $XP = X$ & $\boxed{x_1 + x_2 = 1}$ — (3)

$$[x_1 \;\; x_2] \begin{bmatrix} \frac{1}{4} & 3/4 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = [x_1 \;\; x_2]$$

$$\frac{x_1}{4} + \frac{x_2}{2} = x_1 \Rightarrow -\frac{3}{4}x_1 + \frac{x_2}{2} = 0 \quad \text{— (1)}$$

$$\frac{3x_1}{4} + \frac{x_2}{2} = x_2 \qquad \frac{3x_1}{4} - \frac{x_2}{2} = 0 \Rightarrow -\frac{3x_1}{4} + \frac{x_2}{2} = 0 \quad \text{(2)}$$

eqn's (1) & (2) same.

$$\therefore \; -\frac{3x_1}{4} + \frac{x_2}{2} = 0 \Rightarrow -3x_1 + 2x_2 = 0$$

$$x_1 + x_2 = 1 \qquad \underline{2x_1 + \cancel{2x_2} = 2}$$

$$-5x_1 = -2$$

$$x_1 = 2/5$$

$$\therefore \; \frac{2}{5} + x_2 = 1$$

$$x_2 = 1 - \frac{2}{5} = \frac{3}{5}$$

$$\therefore \; [x_1 \;\; x_2] = [2/5, 3/5]$$

(5)

(5) Recurrent State:- The state is said to be recurrent, if any time that we leave that state we will return to that state in the future with probability one.

Long Answer questions

(1) Define classification of states.

Ans:- **Classification of states:-**

① **Absorbing state:-** If $P_{ii}=1$ then 'i' is said to be absorbing state.

Ex:- $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$

Here absorbing states are 1, 3, because $P_{11}=1$, $P_{33}=1$

② **Transient state:-** If $P_{ii}<1$ then 'i' is said to be transient state

Ex:- $P = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{4} & \frac{3}{4} \end{bmatrix}$

∴ Here transient states are 1, 3. because $P_{11}<1$, $P_{33}<1$

③ **Return state:-** If $P_{ii}^{(n)}>0$ for some 'n' then 'i' is called return state.

∴ Ex:- $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$

∴ Here Return states are 1, 2, 3

because $P_{11}>0$, $P_{22}>0$, $P_{33}>0$

Irreducible State:- If $P_{ij}^{(n)} > 0$ for some 'n' then it is irreducible State

$$A = \begin{bmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

$P_{12}^{(1)} > 0$, $P_{21}^{(1)} > 0$, $P_{22}^{(1)} > 0$ But $P_{11} \not> 0$

$$P^2 = \begin{bmatrix} 1/3 & 2/3 \\ 2/9 & 7/9 \end{bmatrix}$$

$\therefore$ $P_{11} > 0$

$\therefore$ It is irreduciable State

Periodic State:- The Periodic of return state is defined as the GCD of all n Such that

$P_{ij}^{(n)} > 0$, $d_i = GCD\{n, P_{ii}^{(n)} > 0$

if $d_i > 1$ then State 'i' is Called Periodic State

Ex:- $P = \begin{bmatrix} 0 & 1 \\ 1/3 & 2/3 \end{bmatrix}$, $P^2 = \begin{bmatrix} 1/3 & 2/3 \\ 2/9 & 7/9 \end{bmatrix}$

$P_{11}^{(2)} > 0$ $\therefore$ $d_i = 1$

$P_{22}^{(1)} > 0$, $P_{22}^{(2)} > 0$

$\therefore$ $d_i = GCD\{1, 2\} = 1$

$\therefore$ $i = 1, 2$ are aperiodic States.

(2) The transition probability matrix of a Markov Chain is given by $\begin{bmatrix} 0.3 & 0.7 & 0 \\ 0.1 & 0.4 & 0.5 \\ 0 & 0.2 & 0.8 \end{bmatrix}$. Is this matrix irreducible?

Sol:-

Given that $P = \begin{bmatrix} 0.3 & 0.7 & 0 \\ 0.1 & 0.4 & 0.5 \\ 0 & 0.2 & 0.8 \end{bmatrix}$

Here $P_{11}^{(1)} > 0, \quad P_{12}^{(1)} > 0, \quad P_{21}^{(1)} > 0, \quad P_{22}^{(1)} > 0$

$P_{23}^{(1)} > 0, \quad P_{32}^{(1)} > 0, \quad P_{33}^{(1)} > 0$

But $P_{13} = 0, \quad P_{31} = 0$

Then $P^2 = \begin{bmatrix} 0.16 & 0.49 & 0.35 \\ 0.07 & 0.35 & 0.6 \\ 0.02 & 0.24 & 0.74 \end{bmatrix}$

Here $P_{13}^{(2)} > 0, \quad P_{31}^{(2)} > 0$

. So it is irreducible

(3) A fair die tossed repeatedly. If X_n denotes the maximum of the number occuring in the first n tosses, find the transition probability matrix. Find also P^2.

Sol:-

State space $= \{1, 2, 3, 4, 5, 6\}$

Let $X_n =$ max of the number occuring in the first n trials $= 3$ (say)

Then $x_{n+1} = 3$, if $(n+1)^{th}$ trial results is 1, 2, or 3

$= 4$, if $(n+1)^{th}$ trial results is 4

$= 5$, if $(n+1)^{th}$ trial results is 5

$= 6$, if $(n+1)^{th}$ trial results is 6.

$\therefore$ $P\{X_{n+1} = 3 / x_n = 3\} = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 3/6$

$P\{X_{n+1} = i / x_n = 3) = 1/6$ when $i = 4, 5, 6$

$\therefore$ The tpm of Chain is

$$P = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & \frac{2}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 3/6 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 4/6 & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 5/6 & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P^2 = \begin{bmatrix} 1 & 3 & 5 & 7 & 9 & 11 \\ 0 & 4 & 5 & 7 & 9 & 11 \\ 0 & 0 & 9 & 7 & 9 & 11 \\ 0 & 0 & 0 & 16 & 9 & 11 \\ 0 & 0 & 0 & 0 & 25 & 11 \\ 0 & 0 & 0 & 0 & 0 & 36 \end{bmatrix}$$

(4) If the transition probability matrix is $\begin{bmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$

and the initial probabilities are $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ then find the probabilities after three periods. (b) Equilibrium Vector

Sol: G/T $P = \begin{bmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$

$$P_0 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

After one period $P_1 = P_0 . P$

$$= \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 0.25 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

$$= \begin{bmatrix} 0.42 & 0.17 & 0.42 \end{bmatrix}$$

After two periods. $P_2 = P_1 . P$

$$= \begin{bmatrix} 0.42 & 0.17 & 0.45 \end{bmatrix} \begin{bmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

$$= \begin{bmatrix} 0.4 & 0.21 & 0.4 \end{bmatrix}$$

After three Periods $P_3 = P_2 . P$

$$= \begin{bmatrix} 0.4 & 0.21 & 0.4 \end{bmatrix} \begin{bmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

$$= \begin{bmatrix} 0.405 & 0.2 & 0.405 \end{bmatrix}$$

Equilibrium vector :-

let $X = [x_1 \; x_2 \; x_3]$ then $X = XP$ and

$x_1 + x_2 + x_3 = 1$

$$\therefore \; [x_1 \; x_2 \; x_3] = [x_1 \; x_2 \; x_3] \begin{bmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

$$x_1 = 0.5x_1 + 0.5x_2 + 0.25x_3$$
$$x_2 = 0.25x_1 + 0.25x_3$$
$$x_3 = 0.25x_1 + 0.5x_2 + 0.5x_3$$

$$\Rightarrow \quad -0.5x_1 + 0.5x_2 + 0.25x_3 = 0 \quad \text{—} (1)$$
$$0.25x_1 - x_2 + 0.25x_3 = 0 \quad \text{—} (2)$$
$$0.25x_1 + 0.5x_2 - 0.5x_3 = 0 \quad \text{—} (3)$$
$$\& \quad x_1 + x_2 + x_3 = 1 \quad \text{—} (4)$$

Sub $x_1 = 1 - x_2 - x_3$ in (1) & (2)

$$\therefore \; -0.5(1 - x_2 - x_3) + 0.5x_2 + 0.25x_3 = 0$$

$$(1) \Rightarrow \; -0.5 + x_2 + 0.75x_3 = 0 \Rightarrow x_2 + 0.75x_3 = 0.5 \quad \text{—} (5)$$

$$(2) \Rightarrow \; 0.25(1 - x_2 - x_3) - x_2 + 0.25x_3 = 0$$

$$0.25 - 1.25x_2 \cancel{+0.25x_3} \cancel{+0.25x_3} = 0$$

$$\therefore \; \boxed{x_2 = \frac{0.25}{1.25} = \frac{1}{5}}$$

$$(5) \Rightarrow \; \frac{1}{5} + 0.75x_3 = 0.5 \Rightarrow \boxed{x_3 = 2/5}$$

$$(1) \Rightarrow \; -0.5x_1 + \frac{0.5}{5} + 0.25(2/5) = 0$$

$x_1 = 2/5$

(5) Suppose there are two market products of brands A and B respectually. let each of these two brands have exactly 50% of the total market in Same Period and let the market be of a fixed size. The transition matrix is given as follows

	A	B
A	0.9	0.1
B	0.5	0.5

If the initial market share breakdown is 50% for each brand, then determine their market shares in the steady state.

Sol:

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix}$$

The Steady state Vector is

$X = XP$ where $x_1 + x_2 = 1$

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix}$$

$x_1 = 0.9x_1 + 0.5x_2$

$x_2 = 0.1x_1 + 0.5x_2$

$\Rightarrow \left.\begin{array}{l} -0.1x_1 + 0.5x_2 = 0 \\ 0.1x_1 - 0.5x_2 = 0 \end{array}\right\}$ These two eqn's are same

$$\therefore\ 0.1x_1 - 0.5x_2 = 0 \Rightarrow x_1 = \frac{0.5}{0.1}x_2$$

$$x_1 + x_2 = 1$$

$$\frac{0.5}{0.1}x_2 + x_2 = 1$$

$$6x_2 = 1$$

$$x_2 = 1/6$$

$$x_1 = 5/6$$

$$\therefore \text{ Vector is } [x_1 \ \ x_2] = \left[\frac{5}{6} \ \ 1/6\right].$$

(7) The tPm of a markov chain $\{x_n\}$, $n = 1, 2, 3 \ldots$ having 3 states 1, 2 & 3 $P = \begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$ & the initial distribution is $P_0 = (0.7 \ \ 0.2 \ \ 0.1)$ Find (1) $P(x_2 = 3)$ (2) $P\{x_3 = 2, x_2 = 3, x_1 = 3, x_0 = 2\}$.

Sol:- Given that $P = \begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$

$$P^2 = \begin{bmatrix} 0.43 & 0.31 & 0.26 \\ 0.24 & 0.42 & 0.39 \\ 0.36 & 0.35 & 0.29 \end{bmatrix}$$

① $P(x_2=3) = \sum_{i=1}^{3} P\{x_2=3/x_0=i\} \cdot P[x_0=i]$

$= P\{x_2=3|x_0=1] \cdot P(x_0=1) + P\{x_2=3/x_0=2\}$

$\cdot P\{x_0=2\} + P\{x_2=3/x_0=3\} \cdot P\{x_0=3\}$

$= P_{13}^{(2)} \cdot (0.7] + P_{23}^{(2)}(0.2) + P_{33}^{(2)}(0.1)$

$= (0.26)(0.7) + (0.34)(0.2) + 0.29(0.1)$

$= 0.279$

② $P\{x_1=3, x_0=2\} = P\{x_1=3/x_0=2\} \cdot P\{x_0=2\}$

$= P_{23}^{(1)} \times 0.2$

$= 0.2 \times 0.2 = 0.04$

$P\{x_2=3, x_1=3, x_0=2\} = P\{x_2=3/x_1=3, x_0=2\}$
$\times P\{x_1=3, x_0=2\}$

$= P\{x_2=3/x_1=3\} \times P\{x_1=3, x_0=2\}$

$= P_{33}^{(1)} \cdot \times 0.04$

$= 0.3 \times 0.04$

$= 0.012$

$P\{x_3=2, x_2=3, x_1=3, x_0=2\}$

$= P\{x_3=2/x_2=3, x_1=3, x_0=2\}$
$\times P\{x_2=3, x_1=3, x_0=2\}$

$= P\{x_3=2/x_2=3\} \times 0.012$

$= P_{32}^{(1)} \times 0.012 = 0.4 \times 0.012$
$= 0.0048$

UNIT-V

Stochastic process (or) Random process:-

A set of Random variable values $\{x_t\}$ (or) $\{x_1\}$ depending on some real parameters (time t, temprature etc..) is known as stochastic process.

states:-

The values assumed by R.V.

state space:-

The set of all possible values of any individual no. of random process is called state space.

It is denoted by $\{x_t, t\geq 1\} = I$ (or) S.

Ex: when a fair die is tossed, the no. of sixes is stochastic process

If the parameter set is discrete then the state space is discrete.

If the parameter set has infinite values, the state space is continous.

** Classification of stochastic process:-

1)

R.V \ time: x \ t	Continous	Discrete
continous	continous stochastic process	continous stochastic sequence
Discrete	Discrete stochastic process	Discrete stochastic sequence

Stationary stochastic process:-

If the probability distribution do not depends on the time 't' then the random process is called stationary stochastic process.

Deterministic stochastic process:-

A Random process is called deterministic stochastic process if future values of any sample functions can be predicted from its past observations.

Non-Deterministic

A stochastic process is called non deterministic. if future values of any sample functions can't be predicted from its past observations

Markov process:-

A Random process X_n is called markov process if $P\{X(t_{n+1}) < x_{n+1} / X(t_n) < x_n, X(t_{n-1}) \leq x_{n-1} \ldots,$

$X(t_0) \leq x_0\}$.

$\Rightarrow P\{X(t_{n-1}) < x_{n+1} / X(t_n) < x_n\}$

$\Rightarrow P(X_{n+1} / X_n)$.

$X_{n+1}, X_n, X_{n-1}, X_{n-2}, \ldots X_0$

All states of micro processor

(or)

$P\{X_n = k / X_{n-1} = j, X_{n-2} = j_1, \ldots X_0 = J_{n-1}\}$

$\Rightarrow P\{X_n = k / X_{n-1} = j\}$

$\Rightarrow P_{jk}^{(1)}$

4

7

unit step transition probability:-

The probability $P_{jk}^{(1)}$ is called unit step transition probability,

M-step transition probability:-

$$P\{X_{n+m}=k \mid X_n=j\} = P_{jk}^{(m)}.$$

Homogeneous Markov process:-

~~m-step~~ If the transition probability P_{jk} is independent of 'n', then the markov chain is called Homogeneous markov process:-

Non - Homogeneous

If the transition probability P_{jk} is dependent of 'n' then the step is called markov non - Homogeneous markov process.

probability distribution vector:-

A row or column matrix which consists of the probabilities of occurences of markov process then it is known as probability distribution vector.

if $P_1, P_2, P_3, \ldots P_n$ are probabilitys.

then it is $[P_1, P_2 \ldots P_n]$

Transition probability matrix:-

The transition probabilities P_{jk} satisfies

i) $P_{jk} \geq 0$ i.e, non -ve elements.

ii) $\sum_{k=1}^{n} P_{jk} = 1 \ \forall j$. i.e, each row sum $= 1$

Scanned by CamScanner

$$P = [P_{jk}]_{m\times n} = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & & & \\ P_{m1} & P_{m2} & \cdots & P_{mn} \end{bmatrix}$$

This is called Transition probability matrix.

Stochastic matrix:-

A transition probability matrix is called stochastic matrix if it is square matrix.

Regular matrix:-

A tpm is called Regular matrix if it satisfies

i) stochastic matrix

ii) diagonal element shouldn't equal to 1.

iii) All elements of P^m, $m = 2, 3, \cdots$ are positive.

1) which of the following matrices are stochastic matrix

i) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ii) $\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$ iii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

iv) $\begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$ v) $\begin{bmatrix} 0 & 2 \\ 1/4 & 1/4 \end{bmatrix}$

i) stochastic

ii) stochastic

iii) Not stochastic. It is rectangular matrix

iv) not stochastic. Negative element.

v) Row sum $\neq 1$

$\therefore$ Non stochastic

2) which of the following are regular matrices.

i) $A = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix}$ ii) $B = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$

iii) $C = \begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix}$ iv) $D = \begin{bmatrix} 1/2 & 1/4 & 1 \\ 0 & 1/2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

i) It is not Regular. Since, 1 lies on the diagonal.

ii) $B^2 = B \cdot B = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 1/2 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 1/3 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 3/8 & 3/8 & 0 \end{bmatrix}$

$B^3 = B^2 \cdot B = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 3/8 & 3/8 & 0 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 1/2 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 7/16 & 7/16 & 1/8 \end{bmatrix}$

The elements B_{13}, B_{23} are non-zeroes.

iii)
$C^3 = C^2 \cdot C = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{bmatrix}$

$C^4 = C^3 \cdot C = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 1/4 & 1/2 \\ 1/4 & 1/2 & 1/4 \end{bmatrix}$

$C^5 = C^4 \cdot C = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/8 & 1/2 & 3/8 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$

The all elements of P^5 is Non-zero..

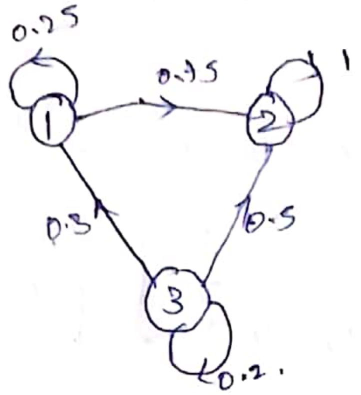
$\therefore$ It is Regular stochastic matrix.

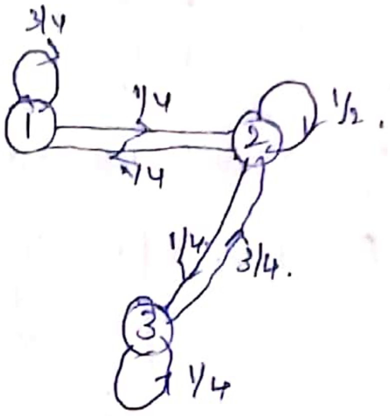
iv) Not a Regular stochastic matrix.

3) Represent the following ~~matrices as a~~ transition matrices as a diagraph.

i) $\begin{bmatrix} 0.25 & 0.75 & 0 \\ 0 & 1 & 0 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$.

ii) $\begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & 3/4 & 1/4 & 0 \\ 2 & 1/4 & 1/2 & 1/4 \\ 3 & 0 & 3/4 & 1/4 \end{array}$





Steady state prob. Distribution :-

If $p^{(0)}$ is initial state probability distribution vector then after 1 step the distribution vector becomes $p^{(1)} = p^{(0)}.P$. Here $P =$ tpm.

After 2 steps:-

$$p^{(2)} = p^{(1)}.P = p^{(0)}.P^2.$$

After n steps:-

$$p^{(n)} = p^{(n-1)}P = p^{(0)}.P^n.$$

If a homogeneous markov chain is regular, every sequence of state probability distribution approaches a unique fixed prob. distribution is called stationary distribution (or) steady state distribution of markov chain.

when, $n \to \infty$

$$\lim_{n \to \infty} p^{(n)} = \pi$$

where $\pi = [\pi_1, \pi_2, -- \pi_n]$

and the steady state distribution satisfies

$\boxed{\pi P = \pi}$ and $\pi_1 + \pi_2 + \dots + \pi_n = 1$

Champman - Kolmogorov's Theorem:-

If P is tpm of homogeneous markov chain then n step tpm ($P^{(n)}$) is equal to P^n.

$$\left[P_{ij}^{(n)}\right] = \left[P_{ij}\right]^n$$

Classification of states:-

i) Irreducible:-

- If $P_{ij}^{(n)} > 0$ for some 'n' and for every i and j, then every state can be reached from every other state.
- i.e., All states are communicate themselves. Then the markov chain is said to be irreducible. The tpm of irreducible chain is called irreducible matrix. otherwise it is reducible.

ii) The state i of markov chain is called return state if $P_{ij}^{(n)} > 0$ for some 'n'.

iii) The period 'd_i' of return state 'i' is defined as the greatest common divisor of all 'm'. there exists $P_{ij}^{(m)} > 0$.

i.e., $GCD\{m, P_{ij}^{(m)} > 0\} = d_i$

iv) The state 'i' is said to be periodic if $d_i > 1$ and aperiodic if $d_i = 1$.

iii) The state 'i' is aperiodic if $p_{ii} \neq 0$

iv) The probability that the chain returns to state i having started from i, for the first time at n^{th} step is denoted by $f_{ii}^{(n)}$, it is called first return time prob, $\{n, f_{ii}^{(n)}, n=1,2,3 \dots\}$ is called the distribution recurrence time state i.

$$F_{ii} = \sum_{i=1}^{\infty} f_{ii}^{(n)}$$

$$\mu_{ii} = \sum_{i=1}^{\infty} n \cdot f_{ii}^{(n)}$$

if $F_{ii}=1$, then the state i is called recurrent (or) persistent ~~of~~ return to the state i it is also called certain.

If $F_{ii}<1$, then the state i is called transient or uncertain.

μ_{ii} is called mean recurrence time of state i.

If μ_{ii} is finite then it is called non-null persistent.

If μ_{ii} is infinite then it is called null persistent.

v) A non null persistant and aperiodic state is called ergodic.

vi) If a markov chain is finite irreducible then all its states are non-null persistant.

vii) $P(X_3=a, X_2=b, X_1=c, X_0=a) = P(X_3=a/X_2=b) P(X_2=b/X_1=c) \cdot P(X_1=c/X_0=a) \cdot P(X_0=a)$

Scanned by CamScanner

1) A Mark

1) If the tpm of a markov chain is $\begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$ then find steady state distribution of chain.

Let $\pi = [\pi_1 \ \pi_2]$ be steady state distr. vector.

$\pi P = \pi$, $\boxed{\pi_1 + \pi_2 = 1} \to ①$

$[\pi_1 \ \pi_2] \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} = [\pi_1 \ \pi_2]$

$\left[\tfrac{1}{2}\pi_2 \quad \pi_1 + \tfrac{1}{2}\pi_2\right] = [\pi_1 \ \pi_2]$

compare same position element.

$\tfrac{1}{2}\pi_2 = \pi_1$, $\pi_1 + \tfrac{1}{2}\pi_2 = \pi_2$.

$\pi_2 = 2\pi_1 \to ②$ $\qquad 2\pi_1 + \pi_2 = 2\pi_2 \to ③$.

from ②, ①.

$\pi_1 + 2\pi_1 = 1$

$3\pi_1 = 1$

$\pi_1 = 1/3$

from ②, $\pi_2 = 2/3$.

steady state dist - $\pi = [1/3 \ \ 2/3]$.

2) If the initial state prob. distribution of a markov chain is $p^{(0)} = \left(\frac{5}{6} \ \frac{1}{6}\right)$ and the tpm of chain is $\begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$ then find the prob. dist. of chain after 2 steps.

$p^{(0)} = \left(\frac{5}{6}, \frac{1}{6}\right)$.

$P = \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}$.

after 2 steps

$p^{(1)} = p^{(0)} \cdot P = \left(\frac{5}{6} \ \frac{1}{6}\right) \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}$

$p^{(1)} = \left(\frac{1}{12} \ \ \frac{11}{12}\right)$

$p^{(2)} = p^{(1)} \cdot P = \left(\frac{1}{12} \; \frac{11}{12}\right)\begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}$

$p^{(2)} = \left(\frac{11}{24} \; \frac{13}{24}\right)$

Note :-

The steady state distribution. is also called as limiting probabilities, probability in the long run invarient probabilities, stationary probabilities, fraction, proportion, How often.

2) A students study habits are as follows. If he studies one night, he is 70% sure not to study next night. On the other hand the prob. that he does not study two nights in the succession is 0.6. In the long run, how often does he study.

S - study at night

T - Not study at night.

$$P = \begin{array}{c|cc} & S & T \\ \hline S & 0.3 & 0.7 \\ T & 0.4 & 0.6 \end{array} \quad \text{Today}$$

Let $\pi = [\pi_1 \;\; \pi_2]$ be steady state dist,

$\pi P = \pi$ where $\pi_1 + \pi_2 = 1 \rightarrow (1)$

$$[\pi_1 \;\; \pi_2] \begin{bmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{bmatrix} = [\pi_1 \;\; \pi_2]$$

$$[0.3\pi_1 + 0.4\pi_2 \quad 0.7\pi_1 + 0.6\pi_2] = [\pi_1 \;\; \pi_2]$$

compare same position element.

$0.3\pi_1 + 0.4\pi_2 = \pi_1$ $\qquad$ $0.7\pi_1 + 0.6\pi_2 = \pi_2$

$-0.7\pi_1 + 0.4\pi_2 = 0$, $\qquad$ $0.7\pi_1 - 0.4\pi_2 = 0$

$-7\pi_1 + 4\pi_2 = 0 \rightarrow (2)$ $\qquad$ $7\pi_1 - 4\pi_2 = 0 \rightarrow (3)$

from ①, ②

② $\Rightarrow 7\pi_1 = 4\pi_2 \Rightarrow \pi_1 = 4/7\, \pi_2$.

① $\Rightarrow \pi_1 + \pi_2 = 1$

$4/7\, \pi_2 + \pi_2 = 1$

$4\pi_2 + 7\pi_2 = 7$

$11\pi_2 = 7$.

$\pi_2 = 7/11$

$\therefore \pi_1 = \frac{4}{7}\pi_2 = \frac{4}{7} \times \frac{7}{11} = \frac{4}{11}$

$\therefore \pi = \left[\frac{4}{11} \quad \frac{7}{11}\right]$.

3) ~~The tpm of~~

3) Two boys B_1, B_2 and Two girls G_1, G_2 are throwing a ball from one to another each boys throws the ball to other boy with prob. 1/2. and to each girl with prob 1/4. On the other hand each girl throws the ball to each boy with prob. 1/2 and never to other girl. In the long run how often does each girl recieved the ball.

Boy → other Boy prob 1/2
→ each girl prob 1/4,

Girl → each boy prob. 1/2.
→ never to other girl.

$$P = \begin{array}{c} \\ B_1 \\ B_2 \\ G_1 \\ G_2 \end{array} \begin{array}{c} \begin{array}{cccc} B_1 & B_2 & G_1 & G_2 \end{array} \\ \begin{bmatrix} 0 & 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/4 & 1/4 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix}_{4\times 4} \end{array}.$$

Let $\pi = [\pi_1 \; \pi_2 \; \pi_3 \; \pi_4]$ be steady state dist. vector

$\pi P = \pi$

& $\pi_1+\pi_2+\pi_3+\pi_4 \to ①$

$$\begin{bmatrix}\pi_1 & \pi_2 & \pi_3 & \pi_4\end{bmatrix}\begin{bmatrix}0 & 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/4 & 1/4 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0\end{bmatrix} = \begin{bmatrix}\pi_1 & \pi_2 & \pi_3 & \pi_4\end{bmatrix}$$

$$\begin{bmatrix}\frac{\pi_2}{2}+\frac{\pi_3}{2}+\frac{\pi_4}{2}, & \frac{\pi_1}{2}+\frac{\pi_3}{2}+\frac{\pi_4}{2} & \frac{\pi_1}{4}+\frac{\pi_2}{4} & \frac{\pi_1}{4}+\frac{\pi_2}{4}\end{bmatrix} = \begin{bmatrix}\pi_1 & \pi_2 & \pi_3 & \pi_4\end{bmatrix}$$

compare same position element.

$\frac{1}{2}(\pi_2+\pi_3+\pi_4)=\pi_1$ $\quad \frac{1}{2}(\pi_1+\pi_3+\pi_4)=\pi_2$ $\quad \frac{1}{4}(\pi_1+\pi_2)=\pi_3$ $\quad \frac{1}{4}(\pi_1+\pi_2)=\pi_4$

$-2\pi_1+\pi_2+\pi_3+\pi_4=0$ $\quad \pi_1-2\pi_2+\pi_3+\pi_4=0$ $\quad \pi_1+\pi_2-4\pi_3=0$ $\quad \pi_1+\pi_2-4\pi_4=0$

$\pi_1=1/3$, $\pi_2=1/3$, $\pi_3=1/6$, $T_4=1/6$.

Steady state distribution :

$$\pi = \begin{bmatrix}\frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6}\end{bmatrix}.$$

4) A Housewife buys 3 kinds of cereals A, B, C. she never buys the same cereal in successive weeks. If she buys cereal A, the next week she buys cereal B. However she buys B (or) C, next week she is 3 times as ~~lightly~~ likely to buy A as other cereal. How often she buys each of the 3 cereals.

A → B.

B → 3 times as likely to buy A as other cereal C

C → " " " " " " " B

$$P = \begin{array}{c|ccc} & A & B & C \\ A & 0 & 1 & 0 \\ B & 3/4 & 0 & 1/4 \\ C & 3/4 & 1/4 & 0\end{array}$$

let $\pi = [\pi_1\ \pi_2\ \pi_3]$ be steady state distribution vector.

$\pi P = \pi$ & $\pi_1 + \pi_2 + \pi_3 = 1 \rightarrow (1)$

$$[\pi_1 \quad \pi_2 \quad \pi_3] \begin{bmatrix} 0 & 1 & 0 \\ 3/4 & 0 & 1/4 \\ 3/4 & 1/4 & 0 \end{bmatrix} = [\pi_1 \quad \pi_2 \quad \pi_3]$$

$$\left[\frac{3\pi_2}{4} + \frac{3\pi_3}{4} \quad \pi_1 + \frac{\pi_3}{4} \quad \frac{\pi_2}{4}\right] = [\pi_1 \quad \pi_2 \quad \pi_3].$$

$\frac{3\pi_2}{4} + \frac{3\pi_3}{4} = \pi_1$ $\quad \pi_1 + \frac{\pi_3}{4} = \pi_2$ $\quad \frac{\pi_2}{4} = \pi_3$.

$-4\pi_1 + 3\pi_2 + 3\pi_3 = 0 \rightarrow (2)$ $\quad 4\pi_1 - 4\pi_2 + \pi_3 = 0 \rightarrow (3)$ $\quad \pi_2 - 4\pi_3 = 0, \rightarrow (4)$

By solving (1), (2), (4),

$\pi_1 = 3/7$ $\quad \pi_1 = 3/7$ $\quad \pi_2 = \frac{16}{35}$ $\quad \pi_3 = \frac{4}{35}$

5) The transition prob. matrix of markov chain $\{X_n\}, n = 1, 2, 3 \ldots$ having 3 states 1, 2 & 3 is

$$P = \begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$$ and the initial distribution $p^{(0)} = (0.7 \quad 0.2 \quad 0.1)$

Find i) prob. $P(X_2 = 3)$. ii) $P(X_3 = 2, X_2 = 3, X_1 = 3, X_0 = 2)$

Given $p^{(0)} = (\overset{1}{0.7} \quad \overset{2}{0.2} \quad \overset{3}{0.1})$

$P(X_0 = 1) = 0.7$ $\quad P(X_0 = 2) = 0.2$ $\quad P(X_0 = 3) = 0.1$

i) $P(X_2 = 3) = P^{(2)}(\text{at state } 3)$

$$p^{(1)} = p^{(0)} \cdot P = [0.7 \quad 0.2 \quad 0.1] \begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$$

$$p^{(1)} = [0.22 \quad 0.43 \quad 0.35]$$

$$p^{(2)} = p^{(1)} \cdot P = [0.22 \quad 0.43 \quad 0.35] \begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$$

$$= [0.385 \quad 0.336 \quad 0.279]$$

$\therefore p^{(2)}(\text{at state } 3) = P(X_2 = 3) = 0.279.$

ii) $P(X_3=2, X_2=3, X_1=3, X_0=2) = P(X_3=2 | X_2=3)\, P(X_2=3 | X_1=3)$
$P(X_1=3 | X_0=2) \cdot P(X_0=2)$

$\Rightarrow p_{32}^{(1)} \cdot p_{33}^{(1)}\, p_{23}^{(1)}\, (0.2)$

$\because P(X_{n+1}=k / X_n=j)$

$\Rightarrow (0.4)(0.3)(0.2)(0.2)$

$\Rightarrow 3/625.$

6) 3 Boys A, B, C are throwing a ball to each other. A always throws the ball to B and B always throws ball to C but C is just as likely to throw the ball to B as to A. Show that the process is markovian. Find the transition prob-matrix and classify the states.

$A \to B$

$B \to C$

$C \to B$
$\ \ \to A$

$$P = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \end{array}$$

future values depends on present values.
$\therefore$ The chain is markovian.

Irreducible :-

$$P^2 = P \cdot P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

$$P^2 = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

$$P^3 = P^2 \cdot P = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$$

$$P^4 = P^3 \cdot 3 = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 1/2 & 1/2 \\ 1/4 & 1/2 & 1/4 \end{bmatrix}.$$

$$P^5 = P^4 \cdot P = \begin{bmatrix} 1/4 & 1/4 & 1/2 \\ 1/4 & 1/2 & 1/4 \\ 1/8 & 3/8 & 1/2 \end{bmatrix}$$

$\therefore P_{11}^{(3)} > 0,\ P_{13}^{(2)} > 0,\ P_{21}^{(2)} > 0,\ P_{22}^{(2)} > 0,\ P_{33}^{(2)} > 0.$

$\therefore P_{ij}^{(n)} > 0$ for some 'n'.

$\therefore$ All states are irreducible.

$\therefore$ Matrix is irreducible.

periodic :-

periodicity of state 'A' :-

$P_{11}^{(3)} > 0,\ P_{11}^{(5)} > 0,\ P_{11}^{(6)} > 0, \ldots$

G.C.D $= \{3, 5, 6 \ldots\} = 1 = d_i$.

if $d_i = 1$ then state 'A' is aperiodic.

periodic of B.

$P_{22}^{(2)} > 0,\ P_{22}^{(3)} > 0,\ P_{22}^{(4)}, \ldots$

G.C.D $= \{2, 3, 4 \ldots\} = 1$

if period $(d_1) = 1$

$\therefore$ state B is aperiodic

$\therefore$ all states are aperiodic.

~~state~~ periodic of C

$P_{33}^{(2)} > 0,\ P_{33}^{(3)} > 0,\ P_{33}^{(4)}, \ldots$

G.C.D $= \{2, 3, 4, \ldots\} = 1$

if period $(d_i) = 1$

$\therefore$ state C is aperiodic

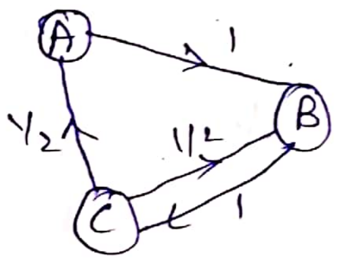
$\therefore$ All states are aperiodic

It has finite states.
The finite irreducible matrix becomes non-null, persistent.

All states are non-null, persistent & aperiodic.
∴ All states are ergodic.

(or)

Diagraph:-



The state A is reachable to state B & C
The state B " " " " C & A
" " C " " " " A & B.

All states are reachable from all other states
∴ The chain is irreducible.
It has infinite states.
∴ finite irreducible matrix becomes non-null persistent.

period of A:-

G.C.D of {3, 5, 7 ...} = 1.
∴ state A is aperiodic

period of B:-

G.C.D {2, 3, 4, 5} = 1
∴ state B is aperiodic

period of C:-

G.C.D = {2, 3, 4, 5, ...} = 1
state C is aperiodic

Note :-

1) Absorbing state :-

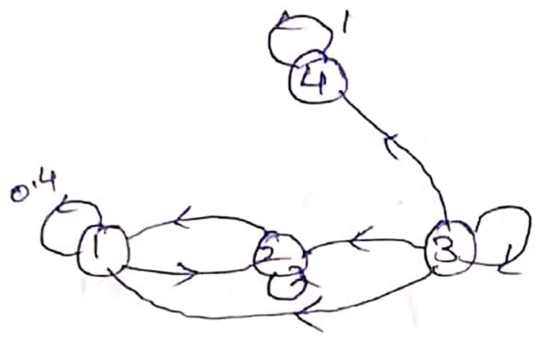
If a state i is called absorbing state if $P_{ii} = 1$.

2) If a matrix with absorbing state then it is not irreducible.

1) Construct the markov chain with transition prob matrix

$$P = \begin{array}{c} A \\ B \\ C \\ D \end{array} \begin{bmatrix} 0.4 & 0.6 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 \\ 0.2 & 0.4 & 0.1 & 0.3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Is this matrix irreducible.



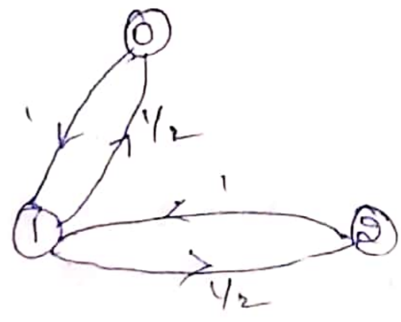
The state 4 is not reachable to state ①,②,③.

∴ The state 4 is absorbing state.

2) Find the nature of the states of the markov chain with the tpm

$$\begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{bmatrix}$$

diagraph :-



The state '0' is reachable from the states 1 & 2

The state '1' is reachable from the states 0 & 2,

The " '2' " " " " 1 & 0

every state is reachable from all other states.

∴ The chain is irreducible.

∴ the chain is finite.

periodicity:-

period of state '0' :- G.C.D $\{2, 4, 6, 8, \ldots\} = 2 = d_i$

period of '0' is $= d_i = 2$.

period of state '1' :-

G.C.D $= \{2, 4, 6, \ldots\} = 2 = d_i$

period of state '1' $= d_i = 2$.

period of state '2' :-

G.C.D $= \{2, 4, 6, \ldots\} = 2 = d_i$.

∴ period of state '2' is $= 2 = d_i$.

∴ All states are non-null, persistent & aperiodic

∴ All states are ergodic.

3) Suppose that the probability of dry day following rainy day is 1/3 and the rainy day following is 1/2. Given that may 1st is dry day. Find that may 3rd is dry day and also may 5th is dry day.

Let D → Dry day

R → Rainy day.

$$P = \begin{array}{c c} & \begin{array}{cc} D & R \end{array} \\ \begin{array}{c} D \\ R \end{array} & \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix} \end{array}$$

Given may 1st is dry day

$$p^{(0)} = \begin{array}{c} \begin{array}{cc} D & R \end{array} \\ \begin{bmatrix} 1 & 0 \end{bmatrix} \end{array}$$

$$p^{(1)} = p^{(0)} \cdot P = [1 \quad 0] \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix}.$$

$p^{(1)} = \begin{bmatrix} 1/2 & 1/2 \end{bmatrix}$

$p^{(2)} = p^{(1)} \cdot P = \begin{bmatrix} 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix}$.

$p^{(2)} = \begin{bmatrix} \frac{1}{4} + \frac{1}{6} & \frac{1}{4} + \frac{2}{6} \end{bmatrix} = \begin{bmatrix} \frac{5}{12} & \frac{7}{12} \end{bmatrix}$

$p^{(3)} = p^{(2)} \cdot P = \begin{bmatrix} \frac{5}{12} & \frac{7}{12} \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix}$

$p^{(3)} = \begin{bmatrix} 0.402 & 0.59 \end{bmatrix}$

prob. of may 3rd dry day $= 0.402$.

$p^{(4)} = p^{(3)} \cdot P$.

$p^{(5)} = p^{(4)} \cdot P$.

4) A man either drives a car or catches the train to go to the office on each day. He never goes two days in a row by train but if he drives one day, the next day he is just as likely to drive again as he ie to travel by train. Now suppose that, on the 1st day of week the man tossed a fair die and drove to if 6 appear. Find

i) prob. that He takes on the 3rd day.

ii) prob. that he drives to work in the long run.

Let $C \rightarrow$ car

$T \rightarrow$ Train

$$P = \begin{array}{c} \\ T \\ C \end{array} \begin{array}{c} \begin{array}{cc} T & C \end{array} \\ \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} \end{array}$$

He tossed a fair die i.e if it shows '6' then he goes by car.

$$p^{(1)} = \begin{array}{c} \begin{array}{cc} T & C \end{array} \\ \begin{bmatrix} \frac{5}{6} & \frac{1}{6} \end{bmatrix} \end{array}$$

i) prob. of 3rd day by Train.

$$P^{(2)} = P^{(1)} \cdot P = \begin{bmatrix} \frac{5}{6} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{12} & \frac{5}{6} + \frac{1}{12} \end{bmatrix} = \begin{bmatrix} \frac{1}{12} & \frac{11}{12} \end{bmatrix}$$

$$P^{(3)} = P^{(2)} \cdot P = \begin{bmatrix} \frac{1}{12} & \frac{11}{12} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{11}{24} & \frac{1}{12} + \frac{11}{24} \end{bmatrix} = \begin{bmatrix} \frac{11}{24} & \frac{13}{24} \end{bmatrix}$$

prob. of 3rd day goes by Train is $= 11/24$.