PROBABILITY AND STATISTICS. 9/4/19

UNIT-1:

Random Variables

→ Discrete and continuous random variables - Properties - Expectation of discrete and continuous random variables and variancy of a sum.

UNIT-2

Probability Distributions.

→ Binomial - Poison and normal - Evaluation of statistical parameters for these three distributions and problems.

UNIT-3

Basic statistics.

→ Corelation and regression - Rank corelation - curve fitting by method of least squares - Fitting of straight lines - second degree parabolas (polynomials) - power and exponential curves.

UNIT-4

Testing of Hypothesis - I

→ Null and alternative hypothesis - critical region and types of errors - Test of significance - Large sample test for single proportion - Difference of

Scanned with CamScanner

proportions – single mean – Difference of means.

UNIT – 5 :

Testing of Hypothesis – II

→ small samples test for single mean – Difference of two means – Test for ratio of variances (f – test) – chi square test for goodness of fit and independent of attributes.

Scanned with CamScanner

I- RANDOM VARIABLES

12/7/19

INTRODUCTION

→ Probability or a chance is a word we often encounter in our day to day life

→ The branch of mathematics which studies the influence of chance in the theory of probability.

EXPERIMENT:

An experiment is any physical action that is observed and results noted.

For example, Tossing a coin
Rolling a die.

RANDOM EXPERIMENT

An experiments repeated under the same circumstances is called Random experiment

ex: Tossing a coin 10 or 12 times
Rolling a die 5 or 6 times

TRIAL:

Each performance in a random experiment is trial.

OUTCOME

The results of a trial in random experiment is called Outcome.

Ex:

Scanned with CamScanner

SAMPLE SPACE :

Collection of all possible outcomes of a random experiment is called as sample space. Denoted by 'S'

Ex: Tossing a coin

$$S = \{H, T\}$$

→ Tossing a coin 2 times

$$S = \{TT, HH, HT, TH\}$$

EVENT :

Every non empty subset of a sample space is called an event. It is denoted by E

For example; Tossing a coin

$$S = \{T, H\}$$

$$E_1 = \{T\}, E_2 = \{H\}$$

EXHAUSTIVE EVENTS

The total number of possible events in any trial of random experiment is known as exhaustive event.

NOTE

→ The events $E_1, E_2, E_3 \ldots\ldots E_n$ are said to be exhaustive events then

$E_1 \cup E_2 \cup E_3 \cup \ldots\ldots \cup E_n$ = sample space(S)

$$\bigcup_{i=1}^{n} E_i = S$$

Scanned with CamScanner

EXCLUSIVE EVENT

Events from a random experiment are said to be mutually exclusive events if the happening of one event preventing the happening of other event.

NOTE

→ Two events E_i and E_j are said to be mutually exclusive events if

$$E_i \cap E_j = \phi$$

Ex: Two coins are tossed.

$$S = \{HH, TT, HT, TH\}$$

$$E_1 = \{HH, TH\}, \; E_2 = \{TT, HT\}$$

$$E_1 \cap E_2 = \phi$$

EQUALLY LIKED EVENTS

Events are said to equally liked events if one cannot be expected performance to the other events. i.e the chance of getting each and event is same.

Ex: The experiment of rolling a die all the six possible events or equally liked events.

i.e $E_1 = E_2 = E_3 = E_4 = E_5 = E_6 = 1/6$

$$E_1 = E_2 = 1/2$$

Scanned with CamScanner

PROBABILITY

If any experiment is performed, n is the no. of possible cases and m is the no. of favourable cases of an event E. Then the probability of an event is defined by

$$P(E) = \frac{m}{n} = \frac{\text{The no. of favourable cases}}{\text{Total no. of possible cases.}}$$

$$P(E) = \frac{n(A)}{n(S)}$$

NOTE

→ If E is an event

→ If probability of E is one then E is called sure event or certain event.

→ If probability of E is zero then E is called impossible event.

$$P(E) = 1, \quad P(E) = 0$$

PROPERTIES

→ If 'S' is a sample space and E is any event in random experiment then

- $P(S) = 1$
- E_1, E_2 are any 2 events in a sample space S then

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

Scanned with CamScanner

→ If E_1, E_2 are mutually exclusive events then P(

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

→ A coin is tossed n times, then the sample space contains 2^n elements.

→ A die is thrown n times or n dies are thrown at a time then the sample space contains 6^n elements.

Example:

1) In a single row throw with 2 die, find the probability of throwing a sum is 10, which is a perfect square.

Given that,

favourable events whose sum is 10 then

$S = \{(5,5), (6,4), (4,6)\}$

Two dies are thrown at a time,

$$6^n = 6^2 = 36$$

$$P(S) = \frac{n(A)}{n(S)} = \frac{3}{36} = \frac{1}{12}$$

favourable events which is a perfect square

$S = \{(2,2), (3,1), (1,3), (4,5), (5,4), (3,6), (6,3)\} = 7$

$$P(S) = \frac{n(A)}{n(S)} = \frac{7}{36}$$

Scanned with CamScanner

INDEPENDENT EVENTS:

Two events are said to be independent if the happening of one event does not depend on the happening of other event. then When two events are independent then

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

NOTE

→ If $A_1, A_2, A_3 \ldots\ldots A_n$ are mutually indpndnt events

$$P(A_1 \cap A_2 \cap A_3 \cap \ldots\ldots \cap A_n) = P(A_1)\, P(A_2) \ldots\ldots P(A_n)$$

RANDOM VARIABLES

So far we discussed about the sample space s as a set of all outcomes of an experiment. WKT outcomes of experiment are the elements of sample space S and they are need not to be numbers. Sometimes, we wish to assign a specific number to each outcome such assignment is called random variable.

(OR)

A Random Variable (X) and sample space S is a function

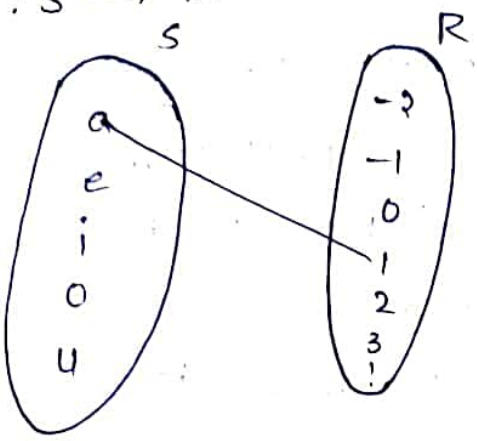
$$X : S \longrightarrow R$$

then assigns a real number $X(s)$ to each sample

Scanned with CamScanner

EXAMPLE

$X : S \longrightarrow R$



→ Random variables are usually classified according to the number of values they are

1) Discript random variables
2) continuous random variables.

1) DISCRIPT RANDOM VARIABLES.

It exist a finite range is called as Discript Random Variables.

Ex: The number of students in a classroom.

$X(x) = \{x : x \text{ is } +ve \text{ integer}\}$

2) CONTINUOUS RANDOM VARIABLES.

If the range of Random variable X is an interval of real numbers then X is an continuous Random variable.

EX: The height of the students in the classroom 4 feets to 6 feets.

DISCRIPT RANDOM VARIABLE PROBABILITY DISTRIBUTION FUNCTION.

let X be a random variable then

Scanned with CamScanner

probability distribution function $F(x)$ is defined by $F(X) = P(X \leq x)$

EXPECTATION OF PROBABILITY DISTRIBUTION FUNCTION.

suppose a random variable X takes the values $x_1, x_2 \ldots x_n$ and the respective probabilities $P_1, P_2 \ldots P_n$. then the expectation X is denoted by $P(X)$ and is defined by $E(x) = \bar{x} = \frac{\Sigma x_i P_i}{\Sigma P_i} = \Sigma x_i P_i$

$$\boxed{E(x) = \bar{x} = \Sigma x_i P_i}$$

*Properties of Expectations.

If X is a random variable and k is any constant

1) $E(aX+b) = aE(X)+k$

2) $E(X+k) = E(X)+k$

3) $E(k) = k$

1) $E(ax+b) = aE(x)+k$

Given that,

X is random variable and k is any constant WKT by the definition of expectation

$$E(X) = \bar{x} = \frac{\Sigma x_i P_i}{\Sigma P_i}$$

$$E(X) = \bar{x} = \Sigma x_i P_i \quad (\Sigma P_i = 1)$$

Scanned with CamScanner

$$E(ax+k) = \bar{x} = \Sigma(ax_i+k)P_i$$

$$= \Sigma(ax_iP_i + P_ik)$$

$$= \Sigma ax_iP_i + \Sigma P_ik$$

$$= a\Sigma x_iP_i + k\Sigma P_i$$

$$\boxed{E(ax+b) = aE(x) + k}$$

2) $E(X+K) = E(X)+K$

$$E(x+k) = \bar{x} = \Sigma(x_i+k)P_i$$

$$= \Sigma(x_iP_i + P_ik)$$

$$= \Sigma x_iP_i + \Sigma P_ik$$

$$= \Sigma x_iP_i + k\Sigma P_i$$

$$E(X+k) = E(X)+k$$

3) $E(K) = K$

$$= \Sigma$$

$$= \bar{x} = \Sigma(x_iP_i)$$

$$= \Sigma x_i$$

$$= k.$$

Scanned with CamScanner

16/7/19

Probability mass function (or) Discrete Probability distribution.

Suppose X is a discrete random variable taking atmost countable no of values $x_1, x_2, x_3 \cdots x_n$. it is with each possible outcome will assosiate a number P_i (i varies from 1 to n) is called probability of mass function of x_i

It must satisfy the following conditions

1) $P_i > 0$

2) $\Sigma P_i = 1$

Then the function P is called probability mass function of random variable.

MEAN

let $x_1, x_2, x_3 \cdots x_n$ are the values axe of random variable X with the corresponding probabilities $P_1, P_2, P_3 \cdots P_n$ the mean of random variable is

$$\bar{x} = \frac{\Sigma x_i P_i}{\Sigma P_i} = \Sigma x_i P_i = E(X)$$

$$\boxed{\text{mean } (\mu) = \bar{x} = \Sigma x_i P_i}$$

'μ' is mean

Scanned with CamScanner

VARIANCE

$$V(x) = V(x) = \sum_{i=1}^{n} (x_i - \bar{x})^2 P_i$$

1) show that $V(x) = E(x^2) - [E(x)]^2$

$$V(x) = \sum_{i=1}^{n} (x_i - \bar{x})^2 P_i$$

$$= \sum_{i=1}^{n} (x_i^2 + \bar{x}^2 - 2x_i\bar{x}) P_i$$

$$= \sum_{i=1}^{n} (P_i x_i^2 + \bar{x}^2 P_i - 2x_i P_i \bar{x})$$

$$= \Sigma x_i^2 P_i + \Sigma \bar{x}^2 P_i - \Sigma 2x_i \bar{x} P_i$$

$$= \Sigma x_i^2 P_i + \bar{x}^2 \Sigma P_i - 2\bar{x} \Sigma x_i P_i \quad (\because \Sigma P_i = 1$$

$$= \Sigma x_i^2 P_i + \bar{x}^2 - 2\bar{x}\bar{x} \qquad \Sigma x_i P_i = \bar{x})$$

By def of expectation

$E(x) = \Sigma x_i P_i$ then $E(x^2) = \Sigma x_i^2 P_i$

$$= E(x^2) - (\bar{x})^2$$

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

NOTE

1) $E(x) = \Sigma x_i P_i$

2) $E(x^2) = \Sigma x_i^2 P_i$

3) $V(x) = \sum_{i=1}^{n} (x_i - \bar{x})^2 P_i$

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

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

Scanned with CamScanner

2) If x is discrete random variable then show that $v(ax+b) = a^2 v(x)$

let $v(x)$ is variance of random variable x ;

let $Y = ax+b$ —— ①

Now applying expectation on both sides

$$E(Y) = E(ax+b)$$

$$E(Y) = aE(x)+b \text{ —— ②}$$

Now, equ ① − eqn ②

$$Y - E(Y) = ax+b - aE(x) - b$$

$$Y - E(Y) = ax - aE(x)$$

$$Y - E(Y) = a(X - E(x))$$

squaring on both sides

$$[Y - E(Y)]^2 = a^2 [X - E(x)]^2$$

Now applying expectation E on B.S

$$E[Y - E(Y)]^2 = a^2 E[x - E(x)]^2$$

$$v(Y) = a^2 v(x) \quad (\because v(x) = E[x - E(x)]^2)$$

$$\boxed{v(ax+b) = a^2 v(x)}$$

**3) A random variable x as the following probability function

Scanned with CamScanner

X	0	1	2	3	4	5	6	7	8
P(X)	$\frac{K}{45}$	$\frac{K}{15}$	$\frac{K}{9}$	$\frac{K}{5}$	$\frac{2K}{45}$	$\frac{6K}{45}$	$\frac{7K}{45}$	$\frac{8K}{45}$	$\frac{4K}{45}$

Then determine i) K ii) Mean iii) Variance

We know that,

$$\sum_{i=0}^{8} P_i = 1$$

(i) $\frac{K}{45} + \frac{K}{15} + \frac{K}{9} + \frac{K}{5} + \frac{2K}{45} + \frac{6K}{45} + \frac{7K}{45} + \frac{8K}{45} + \frac{4K}{45} = 1$

$\frac{45K}{45} = 1 \Rightarrow 45K = 45$

$\boxed{K = 1}$

(ii) Mean

WKT,

$$\bar{x} = \frac{\sum x_i P_i'}{\sum P_i'} = \sum x_i P_i$$

$$\bar{x} = \frac{K}{45} + 2\left(\frac{K}{9}\right) + 3\left(\frac{K}{5}\right) + 4\left(\frac{2K}{45}\right) + 5\left(\frac{6K}{45}\right) + 6\left(\frac{7K}{45}\right) + 7\left(\frac{8K}{45}\right) + 8\left(\frac{4K}{45}\right)$$

$$\bar{x} = \frac{K1}{45} + \frac{2}{9} + \frac{3}{5} + \frac{8}{45} + \frac{30}{45} + \frac{42}{45} + \frac{56}{45} + \frac{32}{45}$$

$$= \frac{1}{15} + \frac{10+27}{45} + \frac{168}{45}$$

Scanned with CamScanner

$$= \frac{208}{45}$$

$$\bar{X} = 4.62$$

(iii) Variance:

$$V(x) = \sum_{i=1}^{n} (x_i - \bar{x})^2 P_i$$

$$V(x) = (0-4.62)^2 \times \frac{1}{45} + (1-4.62)^2 \times \frac{k}{15} + (2-4.62)^2 \times \frac{1}{9} +$$
$$(3-4.62)^2 \times \frac{1}{5} + (4-4.62)^2 \times \frac{2}{45} + (5-4.62)^2 \times \frac{6}{45} +$$
$$(6-4.62)^2 \times \frac{7}{45} + (7-4.62)^2 \times \frac{8}{45} + (8-4.62)^2 \times \frac{4}{45}$$

$$= 0.474 + 0.893 + 0.762 + 0.524 + 0.017 +$$
$$0.019 + 0.296 + 1.007 + 1.015$$

$$V(x) = 4.987$$

Q: A random variable X as the following probability function.

X	0	1	2	3	4	5	6	7
P(x)	0	k	2k	2k	3k	k^2	$2k^2$	$7k^2+k$

Determine;

1) k

2) Evaluate $P(X<6)$, $P(X \geq 6)$, $P(0<X\leq 5)$ and $P(0 \leq X \leq 4)$

3) if $P(X \leq k) > 1/2$ Find minimum value of k

Scanned with CamScanner

(iv) Determine the distribution function of X

(v) mean

(vi) variance.

(i) K

WKT, $\sum_{i=0}^{7} P_i = 1$

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

$9K + 10K^2 = 1$

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

$K = 0.1, -1$

(ii) $P(X < 6)$

$\sum_{i=0}^{5} P_i = 0 + k + 2K + 2K + 3K + K^2$

$= 8K + K^2$

$= 8\left(\frac{1}{10}\right) + \left(\frac{1}{10}\right)^2$

$= \frac{4}{5} + \frac{1}{100}$

$= 0.81$

$P(X \geq 6)$

$\sum_{i=6}^{7} = 2K^2 + 7K^2 + K$

$= 9K^2 + K \Rightarrow \frac{9}{100} + \frac{1}{10}$

$= 0.19$

Scanned with CamScanner

$P(0<X<5)$

$\sum_{x=i}^{4} = k + 2k + 2k + 3k$

$= 8k$

$= 0.8$

$P(0 \leq x \leq 4) = 0 + k + 2k + 2k + 3k$

$= 8k$

$= 0.8$

(iii) $k = 0 \rightarrow P(x=0) = 0 \quad (< 1/2)$

$k = 1 \rightarrow P(x=0) + P(x=1)$

$= 0 + k = \frac{1}{10} = 0.1 \quad (< 1/2)$

$k = 2 \rightarrow P(x=0) + P(x=1) + P(x=2)$

$= 0 + k + 2k = 3k = 0.3 \quad (< 1/2)$

$k = 3 \rightarrow P(x=0) + P(x=1) + P(x=2) + P(x=3)$

$= 0 + k + 2k + 2k$

$= 5k = 0.5 \quad (< 1/2)$

$k = 4 \rightarrow P(x=0) + P(x=1) + P(x=2) + P(x=3) + P(x=4)$

$= 0 + k + 2k + 2k + 3k$

$= 8k = 0.8 \quad (> 1/2)$

$k = 5 \rightarrow P(x=0) + P(x=1) + P(x=2) + P(x=3) +$

$P(x=4) + P(x=5)$

Scanned with CamScanner

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

$= 8K + K^2$

$= \frac{4}{5} + \frac{1}{100}$

$= 0.81$

∴ Minimum value of $K = 4$.

(iv) $K = 0 \rightarrow P(X \leq x)$

$X \quad P(X = 0) = 0$

$K = 1 \rightarrow P(X=0) + P(X=1) = 0.1$

$K = 2 \rightarrow P(X=0) + P(X=1) + P(X=2) = 0.3$

$K = 3 \rightarrow P(X=0) + P(X=1) + P(X=2) + P(X=3) = 0.5$

$K = 4 \rightarrow P(X=0) + P(X=1) + P(X=2) + P(X=3)$

$+ P(X=4) = 0.8$

$K = 5 \Rightarrow P(X \leq 5) = 0.81$

$K = 6 \Rightarrow P(X \leq 6) = 0.83$

$K = 7 \Rightarrow P(X \leq 7) = 1$

(v) Mean

$$\bar{X} = \frac{\Sigma x_i P_i}{\Sigma P_i} = \Sigma x_i P_i$$

$\bar{X} = K + 4K + 6K + 12K + 5K^2 + 10K^2 + 49K^2 + 7K$

$= 30K + 66K^2$

$= \frac{30}{10} + \frac{66}{100}$

$\bar{X} = 3.66$

Scanned with CamScanner

(vi) Variance.

$$V(X) = \sum_{i=1}^{n} (x_i - \bar{x})^2 P_i$$

$$V(X) = (1-3.66)^2 \times \frac{1}{10} + (2-3.66)^2 \times \frac{2}{10} + (3-3.66)^2 \times \frac{2}{10}$$

$$+ (4-3.66)^2 \times \frac{3}{10} + (5-3.66)^2 \times \frac{1}{100} +$$

$$(6-3.66)^2 * \frac{2}{100} + (7-3.66)^2 \times \left(\frac{7}{100} + \frac{11}{10}\right)$$

$$= 1.38048 + 2.02392$$

$$V(X) = 3.40044.$$

19.7.19

A random variable X has the following probability distribution

X	0	1	2	3	4	5	6	7	8
P(x)	a	3a	5a	7a	9a	11a	13a	15a	17a

(i) Find a

(ii) Evaluate $P(X<3)$, $P(X \geq 3)$, $P(0<X<5)$

(iii) Determine the distribution function of X

(i) W.K.T

$$\sum_{i=0}^{8} P_X = 1$$

$$a + 3a + 5a + 7a + 9a + 11a + 13a + 15a + 17a = 1$$

$$81a = 1$$

$$a = 0.012$$

Scanned with CamScanner

(ii) $P(X < 3)$

$= P(X=1) + P(X=2) + P(X=0)$

$= a + 3a + 5a = 9a$

$= 0.111$

iii) $P(X \geq 3)$

$P(X=0) + P(X=1) + P(X=2) + P(X=3)$

$a + 3a + 5a + 7a$

$P(X=3) + P(X=4) + P(X=5) + P(X=6) + P(X=7) + P(X=8)$

$= 7a + 9a + 11a + 13a + 15a + 17a$

$= 72a = \frac{72}{81}$

$= 0.88$

$P(0 < X < 5)$

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

$3a + 5a + 7a + 9a$

$= \frac{24}{81}$

$= 0.296$

(iii) X $P(X \leq x)$

$K = 0 \rightarrow P(X=0) = a = \frac{1}{81} = 0.012$

$K = 1 \rightarrow P(X=0) + P(X=1)$

$= a + 3a = 4a = \frac{4}{81}$

$= 0.049$

Scanned with CamScanner

$K=2 \rightarrow P(X=0)+P(X=1)+P(X=2)$

$= a+3a+5a = 9/81$

$= 0.111$

$K=3 \rightarrow P(X=0)+P(X=1)+P(X=2)+P(X=3)$

$= a+3a+5a+7a = 16/81$

$= 0.197$

$K=4 \rightarrow P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)$

$a+3a+5a+7a+9a$

$= 25/81 = 0.308$

$K=5 \rightarrow P(X=0)+P(X=1)+P(X=2)+P(X=3)+$

$P(X=4)+P(X=5)$

$= a+3a+5a+7a+9a+11a$

$= 36/81 = 0.444$

$K=6 \rightarrow P(X=0)+P(X=1)+P(X=2)+P(X=3)+$

$P(X=4)+P(X=5)+P(X=6)$

$= a+3a+5a+7a+9a+11a+13a$

$= \frac{49}{81} = 0.604$

$K=7 \rightarrow P(X=0)+P(X=1)+P(X=2)+P(X=3)+$

$P(X=4)+P(X=5)+P(X=6)+P(X=7)$

$= a+3a+5a+7a+9a+11a+13a+15a$

$= \frac{64}{81} = 0.709$

Scanned with CamScanner

$K=8 \rightarrow P(X=0)+P(X=1)+P(X=2)+P(X=3)+$
$P(X=4)+P(X=5)+P(X=6)+P(X=7)+P(X=8)$

$= 64a + 17a = 81/81 = 1$

Q: A fair coin is tossed until a head or 5 tails occurs. Find the expected no. of E of tosses of the coin

If head occurs first time there will be only one toss. On the other hand, if first one is tail, second occurs. If head occurs there will be only 2 tosses. suppose second one is also tail third occurs. If head occurs there will be 3 tosses and so on.

$P(1) = P(H) = 1/2$

$P(2) = P(TH) = 1/4$

$P(3) = P(TTH) = 1/8$

$P(4) = P(TTTH) = 1/16$

$P(5) = P(TTTTH) + P(TTTTT)$

$= 1/32 + 1/32 = 1/16$

The probability distribution of x is

X	1	2	3	4	5
P(x)	1/2	1/4	1/8	1/16	1/16

Scanned with CamScanner

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

$= 1\left(\frac{1}{2}\right) + 2(1/4) + 3(1/8) + 4(1/16) + 5(1/16)$

$= 1/2 + 1/2 + 3/8 + 1/4 + 5/16$

$= \frac{31}{16}$

$E(X) = 1.9375$

***Q: A player tosses 3 fair coins. He win Rs 500 if three heads appear, Rs 300 if two heads appear, Rs 100 if one head appear. On the other hand, he losses Rs 1500 if 3 tails occur. Find the expected gain of the player.

Sol: let X denotes the gain.

Then the range of X is

$\{-1500, 100, 300, 500\}$

Then the sample space $S = 2^n = 2^3 = 8$

$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

The probability of all the heads (getting Rs 500) $= P(X=3) = 1/8$

The probability of getting 2 heads (Rs 300) $P(X=2) = 3/8$

The probability of getting 1 head (Rs 100)

$= P(X=1) = 3/8$

Scanned with CamScanner

The probability of getting 3 tails (losing of 1500 Rs) $= P(x=0) = 1/8$

The probability distribution of x is

x	-1500	100	300	500
P(x)	1/8	3/8	3/8	1/8.

$E(x) = \Sigma x_i P_i$

$= (-1500)\frac{1}{8} + (100)\frac{3}{8} + (300)\frac{3}{8} + (500)\frac{1}{8}$

$E(x) = 25$ Rs. 22/7/19

**Q: From a lot of 10 items containing 3 defective items, a sample of 4 items is drawn at random. Let the Random variable x denotes the number of defective items in the sample space. Find the probability distribution of x, when the spample space is drawn without replacement.

The total no. of items = 10

No. of Defective items = 3

No. of good items = 07

let x denotes the no. of defective items in sample space.

$X = \{0, 1, 2, 3\}$

$P(X=0) = P(\text{no defective items})$

Scanned with CamScanner

$= \frac{{}^7C_4}{{}^{10}C_4}$ $\quad {}^nC_r = \frac{n!}{(n-r)!\times r!}$

$= \frac{7!/3!\times 4!}{10!/6!\times 4!}$

$= \frac{7!}{3!} \times \frac{6!}{10!}$

$= \frac{5040}{6} \times \frac{720}{3628800}$

$= 1/6$

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

$= \frac{{}^3C_1 \cdot {}^7C_3}{{}^{10}C_4} = \frac{1}{2}$

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

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

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

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

Probability Distribution of random variable X

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

Scanned with CamScanner

Q: A sample of four items is selected at random from a box containing 12 items of which 5 are detective. Find the expected no. of E of detective items.

The total no. of items = 12.

No. of detective items = 5

No. of good items = 7

let X denotes the no. of detective items in sample space.

$X = \{0, 1, 2, 3, 4\}$

$P(X=0) = P(\text{no detective items})$

$= \frac{7C_4}{12C_4}$

$= \frac{7}{99}$

$P(X=1) = P(1 \text{ detective}, 3 \text{ good})$

$= \frac{5C_1 \cdot 7C_3}{12C_4} = \frac{35}{99}$

$P(X=2) = P(2 \text{ detective}, 2 \text{ good})$

$= \frac{5C_2 \cdot 7C_2}{12C_4} = \frac{14}{33} = \frac{42}{99}$

$P(X=3) = P(3 \text{ detective}, 1 \text{ good})$

$= \frac{5C_3 \cdot 7C_1}{12C_4} = \frac{14}{99}$

Scanned with CamScanner

$$P(X=4) = P(\text{Total defective items})$$

$$= \frac{5C_4}{12C_4} = \frac{1}{99}$$

Probability distribution of random variable

X	0	1	2	3	4
P(X)	7/99	35/99	42/99	14/99	1/99

WKT $E(X) = \Sigma x_i P_i$

$$E(X) = 0 \times \frac{7}{99} + 1 \times \frac{35}{99} + 2 \times \frac{42}{99} + 3 \times \frac{14}{99} + 4 \times \frac{1}{99}$$

$$= \frac{35+84+42+4}{99}$$

$$E(X) = \frac{165}{99} = 1.6$$

Q: Given that $f(x) = \frac{k}{2x}$, is a probability distribution for a random variable X that can take on the values $x = 0, 1, 2, 3, 4$

(i) Find k (ii) Mean (iii) Variance

Given,

$$f(x) = \frac{k}{2x}$$

$x = 0, 1, 2, 3, 4$

Scanned with CamScanner

for $x=0 \rightarrow f(x) = \frac{k}{2x0} = 0$

$x=1 \rightarrow f(1) = \frac{k}{2}$

$x=2 \rightarrow f(2) = \frac{k}{4}$

$x=3 \rightarrow f(3) = \frac{k}{6}$

$x=4 \rightarrow f(4) = \frac{k}{8}$.

Prabability distribution of random variable X

X	0	1	2	3	4
P(X)		K/2	K/4	K/6	K/8

(i) find K

$$\Sigma f_i = \frac{k}{2} + \frac{k}{4} + \frac{k}{6} + \frac{k}{8} = 1$$

$$= \frac{25k}{24} = 1 \Rightarrow k = \frac{24}{25} = 0.96$$

(ii) Mean,

WKT $\Sigma x_i P_i$

$$\Rightarrow \left(1 \times \frac{k}{2}\right) + \left(2 \times \frac{k}{4}\right) + \left(3 \times \frac{k}{6}\right) + \left(4 \times \frac{k}{8}\right)$$

$\Rightarrow 1.92$

(iii) $V(x) = \Sigma (x_i - \bar{x})^2 P_i$

$$V(x) = (1-1.92)^2 \frac{(0.96)}{2} + (2-1.92)^2 \frac{(0.96)}{4} + (3-1.92)^2 \frac{(0.96)}{6} + (4-1.92)^2 \frac{(0.96)}{8}$$

Scanned with CamScanner

$V(x) = 0.406 + (1.536 \times 10^{-3}) + (0.186) + 0.519$

$V(x) = 1.112$

Q: A player 3 fair coins. He win Rs 100/- if head appear, Rs 200/- if 2 heads appear. on the other hand he loss Rs 500/- if no head appears. Determine the expected value of E of the game and is the game favourable to the player?

let x denotes the gain

Then the range of x is

$\{-500, 100, 200\}$

Then the sample space $S = 2^n = 2^2 = 4$

$S = \{HH, HT, TH, TT\}$

The probability of all heads (getting Rs 200) = 1/4

The probability of 1 head (getting Rs 100) = 2/4

The probability of not getting 2 heads (getting Rs 300) losing

$\simeq 3/4 = 1/4$

The probability distribution of x is

X	-500	100	200
P(x)	1/4	2/4	1/4

Scanned with CamScanner

$E(x) = \sum x_i P_i$

$= (-500 \times \frac{1}{4}) + (100 \times \frac{1}{2}) + (200 \times \frac{1}{4})$

$= -25$

∴ The game is not favourable to the player.

Q: Let x denote the minimum of the two number that appear. when a pair of fair die is thrown once, then determine probability distribution function of x

(ii) Expectation

(iii) Variance

Sol: When two dice are thrown, the total no. of outcomes = 36.

The sample space,

$S = \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6)$
$(2,1),(2,2),(2,3),(2,4),(2,5)(2,6)$
$(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)$
$(4,1),(4,2),(4,3),(4,4),(4,5),(4,6)$
$(5,1),(5,2),(5,3),(5,4),(5,5),(5,6)$
$(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}$

let $X = \{1, 2, 3, 4, 5, 6\}$

$P(1) = P(X=1) - (1,1)(1,2)(1,3)(1,4)(1,5)(1,6)$
$(2,1)(3,1)(4,1)(5,1)(6,1)$

Scanned with CamScanner

$= \frac{11}{36}$

$P(2) = P(X=2) - (2,2)(2,3)(2,4)(2,5)(2,6)$
$(3,2)(4,2)(5,2)(6,2)$

$= \frac{9}{36}$

$P(3) = P(X=3) - (3,3)(3,4)(3,5)(3,6)$
$(4,3)(5,3)(6,3)$

$= \frac{7}{36}$

$P(4) = P(X=4) = (4,4)(4,5)(4,6)(5,6)(6,4)$

$= \frac{5}{36}$

$P(5) = P(X=5) = (5,5)(5,6)(6,5)$

$= \frac{3}{36}$

$P(6) = P(X=6) = (6,6)$

$= \frac{1}{36}$

Probability distribution of X

X	1	2	3	4	5	6
P(X)	11/36	9/36	7/36	5/36	3/36	1/36

Scanned with CamScanner

(ii) Expectation

WKT, $E(x) = \Sigma x_i P_i$

$= \left(1 \times \frac{11}{36}\right) + \left(2 \times \frac{9}{36}\right) + \left(3 \times \frac{7}{36}\right) + \left(4 \times \frac{5}{36}\right)$

$+ \left(5 \times \frac{3}{36}\right) + \left(6 \times \frac{1}{36}\right)$

$E(x) = \frac{91}{36} = 2.527.$

(iii) Variance

$V(x) = \Sigma (x_i - \bar{x})^2 P_i$

$V(x) = (1-2.527)^2 \frac{11}{36} + (2-2.529)^2 \times \frac{9}{36} +$

$(3-2.527)^2 \times \frac{7}{36} + (4-2.527)^2 \times \frac{5}{36} +$

$(5-2.527)^2 \times \frac{3}{36} + (6-2.527)^2 \times \frac{1}{36}.$

$= 0.8254 + 1.1460.$

$V(x) = 1.9674$

Q: Two die are thrown. Let x assign to each point (a,b) in S the maximum of its numbers i.e $X(a,b) = \max(a,b)$. Find the probability distribution and also find mean and variance of distribution.

When 2 dice are thrown the total no. of outcomes = 36.

Scanned with CamScanner

$S = \{(1,1)(1,2)(1,3)(1,4)(1,5)(1,6)$
$(2,1)(2,2)(2,3)(2,4)(2,5)(2,6)$
$(3,1)(3,2)(3,3)(3,4)(3,5)(3,6)$
$(4,1)(4,2)(4,3)(4,4)(4,5)(4,6)$
$(5,1)(5,2)(5,3)(5,4)(5,5)(5,6)\}$
$(6,1)(6,2)(6,3)(6,4)(6,5)(6,6)$

let $X = \{1,2,3,4,5,6\}$

$P(1) = P(X=1) = (1,1)$

$= \frac{1}{36}$

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

$= \frac{3}{36}$

$P(3) = P(X=3) = (1,3)(2,3)(3,2)(3,1)(3,3)$

$= \frac{5}{36}$

$P(4) = P(X=4) = (1,4)(2,4)(3,4)(4,4)$
$(4,3)(4,2)(4,1)$

$= \frac{7}{36}$

$P(5) = P(X=5) = (1,5)(2,5)(3,5)(4,5)(5,5)$
$(5,1)(5,2)(5,3)(5,4)$

$= \frac{9}{36}$

$P(6) = P(X=6) = (1,6)(2,6)(3,6)(4,6)(5,6)(6,6)$
$(6,5)(6,4)(6,3)(6,2)(6,1)$

$= \frac{11}{36}$

Scanned with CamScanner

Probability distribution of X

X	1	2	3	4	5	6
P(X)	1/36	3/36	5/36	7/36	9/36	11/36

(ii) Expectation

WKT, $E(x) = \Sigma x_i P_i$

$= \left(1 \times \frac{1}{36}\right) + \left(2 \times \frac{3}{36}\right) + \left(3 \times \frac{5}{36}\right) + \left(4 \times \frac{7}{36}\right) +$

$\left(5 \times \frac{9}{36}\right) + \left(6 \times \frac{1}{36}\right)$

$= 0.027 + 0.166 + 0.416 + 0.777 + 1.25 +$

$0.1666.$

$= 2.802$

(iii) Variance

$v(x) = (x_i - \bar{x})^2 P_i$

$= (1 - 2.802)^2 \times \frac{1}{36} + (2 - 2.802)^2 \frac{3}{36} +$

$(3 - 2.802)^2 \times \frac{5}{36} + (4 - 2.802) \times \frac{7}{36} +$

$(5 - 2.802)^2 \times \frac{9}{36} + (6 - 2.802)^2 \times \frac{11}{36}$

$0.428 + 4.33278$

$v(x) = 4.76078$

Scanned with CamScanner

26/7/19

+++

Q: The random variable X is defined as the sum of the numbers on faces when two dice are thrown. Find mean of X.

Sol let X be the random variable.

The sum of the numbers on the faces when 2 dice are thrown. X is a random variable.

when 2 dice is thrown total outcomes is 36

$X = \{2,3,4,5,6,7,8,9,10,11,12\}$.

$S = \{(1,1)(1,2)(1,3)(1,4)(1,5)(1,6)$
$(2,1)(2,2)(2,3)(2,4)(2,5)(2,6)$
$(3,1)(3,2)(3,3)(3,4)(3,5)(3,6)$
$(4,1)(4,2)(4,3)(4,4)(4,5)(4,6)$
$(5,1)(5,2)(5,3)(5,4)(5,5)(5,6)$
$(6,1)(6,2)(6,3)(6,4)(6,5)(6,6)\}$

$P(2) = P(X=2) \rightarrow (1,1) = 1 = 1/36$

$P(3) = P(X=3) \rightarrow (2,1)(1,2) = 2 = 2/36$

$P(4) = P(X=4) \rightarrow (2,2)(3,1)(1,3) = 3 = 3/36$

$P(5) = P(X=5) \rightarrow (1,4)(4,1)(3,2)(2,3) = 4/36$

$P(6) = P(X=6) \rightarrow (1,5)(5,1)(3,3)(4,2)(2,4)$
$= 5/36$

$P(7) = P(X=7) \rightarrow (1,6)(6,1)(5,2)(2,5)(3,4)$
$(4,3) = 6/36$

Scanned with CamScanner

$P(8) = P(X=8) \rightarrow (4,4)\ (6,2)\ (2,6)\ (4,4)$
$(5,3)\ (3,5) = 5/36$

$P(9) = P(X=9) \rightarrow (3,6)\ (6,3)\ (4,5)\ (5,4) = 4/36$

$P(10) = P(X=10) \rightarrow (6,4)\ (4,6)\ (5,5) = 3/36$

$P(11) = P(X=11) \rightarrow (6,5)\ (5,6) = 2/36$

$P(12) = P(X=12) \rightarrow (6,6) = 1/36.$

Probability distribution of X

X	2	3	4	5	6	7	8	9	10	11	12
P(X)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Mean of $\bar{X} = \sum_{x=2}^{12} x_i P_i$

$= 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + 4 \times \frac{3}{36} + 5 \times \frac{4}{36} + 6 \times \frac{5}{36} +$
$7 \times \frac{6}{36} + 8 \times \frac{5}{36} + 9 \times \frac{4}{36} + 10 \times \frac{3}{36} +$
$11 \times \frac{2}{36} + 12 \times \frac{1}{36}$

$\bar{x} = 7$

30/7/19

CONTINUOUS RANDOM VARIABLES.

→ let, consider the small interval $\left[x - \frac{dx}{2}, x + \frac{dx}{2}\right]$ of length 'dx' around the point x

→ let

Scanned with CamScanner

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

(ii) $f(x) \geq 0 \;\forall\; x \in R$

(iii) $P(a \leq x \leq b) = \int_{a}^{b} f(x)\,dx$

MEAN :

let X be a continuous Random Varia -ble and their probability density function of X over the interval [a,b]. Then mean,

$$\mu = \int_{a}^{b} x\, f(x)\,dx$$

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

MEDIAN

It is the point which divides the entire distribution into two equal parts. In case of continuous distribution, median is the point which divides the total area into 2 equal parts thus x is defined from a to b and M is median then,

$$\int_{a}^{M} f(x)\,dx = \int_{M}^{b} f(x)\,dx = 1/2$$

Solving M we get the median.

Scanned with CamScanner

MODE:

Mode is the value of x for which $f(x)$ is maximum. Mode is thus given by $f'(x) = 0$ and $f''(x) < 0$ for $a < x < b$

VARIANCE

The variance of the distribution is given by

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

Q: A continuous random variable has the probability density function.

$$f(x) = \begin{cases} c\,x(2-x) & \text{if } 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

① Find c, (ii) Mean (iii) Variance

Sol:

WKT,

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

$$\int_{-\infty}^{0} f(x)\,dx + \int_{0}^{2} f(x)\,dx + \int_{2}^{\infty} f(x)\,dx = 1$$

$$0 + \int_{0}^{2} c\,x(2-x) + 0 = 1$$

$$c\int_{0}^{2} x(2-x)\,dx = 1$$

$$c\int_{0}^{2} (2x - x^2)\,dx = 1$$

Scanned with CamScanner

$$c\left[\frac{2x^2}{2} - \frac{x^3}{3}\right]_0^2 = 1$$

$$c\left[x^2 - \frac{x^3}{3}\right] = 1$$

$$c\left(4 - \frac{8}{3}\right) = 1$$

$$\frac{4c}{3} = 1$$

$$c = 3/4$$

(ii) Mean

WKT, $\mu = \int_a^b x\, f(x)\, dx$.

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

$$= 0 + \int_0^2 x \times cx(x-2)\, dx + 0$$

$$= c\int_0^2 (x^3 - 2x^2)\, dx$$

$$= c\int_0^2 x^4 = c\int_0^2 (x^3 - 2x^2)\, dx$$

$$= c\int_0^2 (2x^2 - x^3)\, dx$$

$$= c\left[\frac{2x^3}{3} - \frac{x^4}{4}\right]_0^2$$

$$= \frac{3}{4}\left[\frac{16}{3} - \frac{16}{4}\right]$$

$$\Rightarrow \frac{3}{4}\left[\frac{4}{3}\right]$$

$$\mu = 1$$

Scanned with CamScanner

(iii) Variance

WKT,

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

$$V(x) = \int_{-\infty}^{0} x^2 f(x)\,dx - (\mu)^2 + \int_{0}^{2} x^2 f(x)\,dx - (\mu)^2 + \int_{2}^{-\infty} x^2 f(x)\,dx - (\mu)^2$$

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

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

$$= \int_{0}^{2} x^2 \cdot x(2-x)\,dx - (1)^2$$

$$= c\int_{0}^{2} (2x^3 - x^4)\,dx - 1$$

$$= \frac{3}{4}\left[\frac{2x^4}{4} - \frac{x^5}{5}\right]_0^2 - 1$$

$$= \frac{3}{4}\left[\frac{16}{2} - \frac{32}{5}\right] - 1$$

$$= \frac{3}{4}\left[\frac{8}{5}\right] - 1$$

$$= \frac{24}{20} - 1 \Rightarrow \frac{6}{5} - 1$$

$$V(x) = 1/5$$

Q: For continuous probability function

$$f(x) = \begin{cases} k x^2 e^{-x} & \text{if } x \geq 0 \end{cases}$$

Find k, mean, variance.

Scanned with CamScanner

Sol :

WKT,

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

$$\Rightarrow \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\infty} f(x)\,dx = 1$$

$$0 + \int_{0}^{\infty} f(x)\,dx = 1$$

$$\Rightarrow 0 + \int_{0}^{\infty} k x^2 e^{-x}\,dx = 1$$

$$= k \int_{0}^{\infty} x^2 e^{-x}\,dx = 1$$

$$= k \left[x^2(-e^{-x}) - 2x(e^{-x}) + 2(-e^{-x})\right]_0^{\infty} = 1$$

$$= k \left[-x^2 e^{-x} - 2x e^{-x} - 2e^{-x}\right]_0^{\infty} = 1$$

$$= k \left[-2e^{-(0)}\right] = 1$$

$$= +2k = 1$$

$$k = 1/2.$$

(ii) Mean

WKT,

$$\mu = \int_{a}^{b} x f(x)\,dx$$

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

$$\bar{\mu} = 0 + \int_{0}^{\infty} x k x^2 e^{-x}\,dx$$

$$= k \int_{0}^{\infty} x^3 e^{-x}\,dx$$

$$= \frac{1}{2} \int_{0}^{\infty} x^3 e^{-x}\,dx.$$

Scanned with CamScanner

$$\Rightarrow \frac{1}{2}\left[x^3(-e^{-x}) - 3x^2(e^{-x}) + 6x(-e^{-x}) + 6(e^{-x})\right]_0^{\infty}$$

$$\mu = \frac{1}{2}[0+6] \Rightarrow \mu = \frac{6}{2}$$

$$\mu = 3.$$

(iii) Variance,

W.K.T,

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

$$V(x) = 0 + \int_0^{\infty} x^2 \cdot k\, x^2 e^{-x}\,dx - (3)^2$$

$$= k\int_0^{\infty} x^4 e^{-x}\,dx - (3)^2$$

$$= \frac{1}{2}\left[x^4(-e^{-x}) - 4x^3(e^{-x}) + 12x^2(-e^{-x}) - 24x(e^{-x}) + 24(-e^{-x})\right]_0^{\infty} - 9$$

$$= \frac{1}{2}\left[0 - 24(-e^{-x})\right] - 9.$$

$$V(x) = \frac{1}{2}(24) - 9 \Rightarrow 12 - 9.$$

$$V(x) = 3$$

Q: A continuous random variable has a probability density function.

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

Find, K, mean, variance

(i) WKT

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

Scanned with CamScanner

$$\Rightarrow \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\infty} f(x)\,dx = 1$$

$$= 0 + \int_{0}^{\infty} K x e^{-\alpha x}\,dx = 1$$

$$= K\int_{0}^{\infty} x e^{-\alpha x}\,dx = 1$$

$$K\left[x\left(\frac{-e^{-\alpha x}}{\alpha}\right) - 1\left(\frac{e^{-\alpha x}}{\alpha^2}\right)\right]_0^{\infty} = 1$$

$$= K\left[0 + \frac{1}{\alpha^2}\right] = 1$$

$$K = \alpha^2$$

(ii) WKT

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

$$= \int_{-\infty}^{0} x f(x)\,dx + \int_{0}^{\infty} x f(x)\,dx -$$

$$= 0 + \int_{0}^{\infty} x \cdot K x e^{-\alpha x}\,dx$$

$$= K\int_{0}^{\infty} x^2 e^{-\alpha x}\,dx$$

$$= \alpha^2\left[x^2\left(\frac{-e^{-\alpha x}}{\alpha}\right) - 2x\left(\frac{e^{-\alpha x}}{\alpha^2}\right) + 2\left(\frac{-e^{-\alpha x}}{\alpha^3}\right)\right]_0^{\infty}$$

$$= \alpha^2\left[0 + 2\left(\frac{-e^{-0}}{\alpha^3}\right)\right]$$

$$\mu = \frac{2\alpha^2}{\alpha^3} \Rightarrow \mu = \frac{2}{\alpha}.$$

(iii) WKT,

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

Scanned with CamScanner

$$= \int_{-\infty}^{0} x^2 f(x)\,dx - (\mu)^2 + \int_{0}^{\infty} x^2 f(x)\,dx - (\mu)^2$$

$$V(x) = 0 + \int_0^\infty x^2 \cdot k x e^{-dx}\,dx - \left(\frac{2}{d}\right)^2$$

$$= k \int_0^\infty x^3 e^{-dx}\,dx - \left(\frac{2}{d}\right)^2$$

$$= k\left[x^3\left(\frac{e^{-dx}}{-d}\right) - 3x^2\left(\frac{e^{-dx}}{d^2}\right) + 6x\left(\frac{-e^{-dx}}{d^3}\right) - 6\left(\frac{e^{-dx}}{d^4}\right)\right]_0^\infty - \left(\frac{2}{d}\right)^2$$

$$= d^2\left[0 + 6\left(\frac{e^{-d(0)}}{d^4}\right)\right] - \left(\frac{2}{d}\right)^2$$

$$= d^2\left[\frac{6}{d^4}\right] - \frac{4}{d^2}$$

$$= \frac{6}{d^2} - \frac{4}{d^2}$$

$$V(x) = \frac{2}{d^2}$$

Q: A continuous RV X is defined by

$$f(x) = \begin{cases} \frac{1}{16}(3+x)^2 & \text{if } -3 \le x < -1 \\ \frac{1}{16}(6-2x^2) & \text{if } -1 \le x < 1 \\ \frac{1}{16}(3-x)^2 & \text{if } 1 \le x < 3 \\ 0 & \text{otherwise} \end{cases}$$

(i) Verify that $f(x)$ is density function

(ii) Find mean of X

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

Scanned with CamScanner

$$= \int_{-\infty}^{-3} f(x)\,dx + \int_{-3}^{-1} f(x)\,dx + \int_{-1}^{1} f(x)\,dx +$$
$$\int_{1}^{3} f(x)\,dx + \int_{3}^{\infty} f(x)\,dx = 1$$

$$= 0 + \int_{-3}^{-1} \frac{1}{16}(3+x)^2\,dx + \int_{-1}^{1} \frac{1}{16}(6-2x^2) +$$
$$\int_{1}^{3} \frac{1}{16}(3-x)^2 + 0 = 1$$

$$\Rightarrow \frac{1}{16}\int_{-3}^{-1}(9+x^2+6x)\,dx + \frac{1}{16}\int_{-1}^{1}(6-2x^2)\,dx +$$
$$\frac{1}{16}\int_{1}^{3}(9+x^2-6x)\,dx = 1$$

$$\Rightarrow \frac{1}{16}\left[9x + \frac{x^3}{3} + \frac{6x^2}{2}\right]_{-3}^{-1} + \frac{1}{16}\left[6x - \frac{2x^3}{3}\right]_{-1}^{1}$$
$$+ \frac{1}{16}\left[9x + \frac{x^3}{3} - \frac{6x^2}{2}\right]_{1}^{3} = 1$$

$$\Rightarrow \frac{1}{16}\left[9(-1+3) + \frac{(-1+3)^3}{3} + 3(-1+3)^2\right] +$$
$$\frac{1}{16}\left[6(1+1) - \frac{2}{3}(1+1)^3\right] + \frac{1}{16}\left[9(3-1) +\right.$$
$$\left.\frac{(3-1)^3}{3} - 3(3-1)^2\right] = 1$$

$$\Rightarrow \frac{1}{16}\left[18 + \frac{8}{3} + 12\right] + \frac{1}{16}\left[12 - \frac{16}{3}\right] +$$
$$\frac{1}{16}\left[18 + \frac{8}{3} - 12\right] = 1$$

$$\Rightarrow \frac{49}{24} + \frac{5}{12} + \frac{13}{24} = 1$$

$$\Rightarrow 1$$

Scanned with CamScanner

(11) Mean

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

$$\frac{1}{16}\int_{-3}^{-1} x(3+x)^2\,dx + \frac{1}{16}\int_{-1}^{1} x(6-2x^2)\,dx + \frac{1}{16}\int_{1}^{3} x(3-x)^2\,dx.$$

$$= \frac{1}{16}\int_{-3}^{-1} x(9+x^2+6x)\,dx + 0 + \frac{1}{16}\int_{1}^{3} x(9-6x+x^2)\,dx$$

$$= \frac{1}{16}\left[\frac{9x^2}{2}+\frac{6x^3}{3}+\frac{x^4}{4}\right]_{-3}^{-1} + \frac{1}{16}\left[\frac{9x^2}{3}-\frac{6x^3}{3}+\frac{x^4}{4}\right]_{1}^{3}$$

$$= \frac{1}{16}\left[\left(\frac{9}{2}-2+\frac{1}{4}\right)-\left(\frac{81}{2}-54+\frac{81}{4}\right)\right] + \frac{1}{16}\left[\left(\frac{81}{2}-54+\frac{81}{4}\right)-\left(\frac{9}{2}-2+\frac{1}{4}\right)\right].$$

$= 0$

2/08/19

Q: The probability density function of $f(x)$ of a random variable $f(x) = ce^{-|x|}$, if $-\infty < x < \infty$ show that $c = 1/2$ And find mean and variance of distribution also find the probability that variance lies blw 0 & 4.

Sol: Given that,

$f(x) = ce^{-|x|}$ if $-\infty < x < \infty$

Scanned with CamScanner

WKT,

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

$$\int_{-\infty}^{\infty} c e^{-|x|}\,dx = 1$$

$$c\int_{-\infty}^{\infty} e^{-|x|}\,dx = 1$$

[$\int_{-a}^{a} f(x)\,dx = 0$ if $f(x)$ is odd function

i.e, $f(-x) = -f(x)$

$\int_{-a}^{a} f(x)\,dx = 2\int_{0}^{a} f(x)\,dx$ if $f(x)$ is even

i.e $f(-x) = f(x)$]

$$c\,2\int_{0}^{\infty} e^{-|x|}\,dx = 1 \quad (\because e^{-|x|} \text{ is even function})$$

$$2c\int_{0}^{\infty} e^{-x}\,dx = 1$$

$$2c\left[-e^{-x}\right]_{0}^{\infty} = 1$$

$$2c = 1$$

$$c = 1/2$$

WKT,

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

$$= \int_{-\infty}^{\infty} x \cdot c e^{-|x|}\,dx$$

$$= c\int_{-\infty}^{\infty} x e^{-|x|}\,dx = 0 \quad [\because \text{odd function}]$$

Scanned with CamScanner

WKT,

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

$$= \int_{-\infty}^{\infty} x^2 \cdot c e^{-|x|}\,dx - 0$$

$$= c\int_{-\infty}^{\infty} x^2 e^{-|x|}\,dx$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} 2x^2 e^{-|x|}\,dx.$$

$$= \int_{0}^{\infty} x^2 e^{-|x|}\,dx$$

$$= \left[x^2(-e^{-x}) - 2x(e^{-x}) + 2(-e^{-x})\right]_0^{\infty}$$

$$= \left[a^2(0-1) - 2x(\right]$$

$$\to \left[0 - 2(-e^{-0})\right]$$

$$= 2$$

$$V(x) = 2.$$

$$\int_0^4 f(x)\,dx$$

Q: If the probability density function of a random variable is given by

$$f(x) = \begin{cases} k(1-x^2) & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the value of k and the probabilities that random variable having this probability will take on a value b/w

i) b/w 0.1 & 0.2

ii) greater than 0.5

Scanned with CamScanner

Given that,

$$f(x) = \begin{cases} k(1-x^2) & \text{for } 0<x<1 \\ 0 & \text{otherwise} \end{cases}$$

WKT,

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

$$= \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{\infty} f(x)\,dx = 1$$

$$= 0 + \int_{0}^{1} f(x)\,dx + 0 = 1$$

$$= \int_{0}^{1} k(1-x^2)\,dx = 1$$

$$k\int_{0}^{1} (1-x^2)\,dx = 1$$

$$k\left[x - \frac{x^3}{3}\right]_0^1 = 1$$

$$k\left[1 - \frac{1}{3}\right] = 1$$

$$2k = 3$$

$$k = 3/2$$

→ Probability takes value blw 0.1 & 0.2

$$P(0.1 < x < 0.2)$$

$$\Rightarrow \int_{0.1}^{0.2} f(x)\,dx$$

$$= \int_{0.1}^{0.2} k(1-x^2)\,dx$$

Scanned with CamScanner

$$= \frac{3}{2}\int_{0.1}^{0.2}(1-x^2)\,dx$$

$$= \frac{3}{2}\left[x - \frac{x^3}{3}\right]_{0.1}^{0.2}$$

$$= \frac{3}{2}\left[(0.2-0.1) - \frac{(0.2-0.1)^3}{3}\right]$$

$$= 0.1465.$$

$$P(0.1 < x < 0.2) = 0.1465$$

ii) $P(x > 0.5) = \int_{0.5}^{\infty} f(x)\,dx$

$$= \int_{0.5}^{\infty} k(1-x^2)\,dx$$

$$= k\int_{0.5}^{1}(1-x^2)\,dx + k\int_{1}^{\infty}(1-x^2)\,dx$$

$$= \frac{3}{2}\left[x - \frac{x^3}{3}\right]_{0.5}^{1} + 0$$

$$\left(= \frac{3}{2}\left[(1-0.5) - \frac{(1-0.5)^3}{3}\right]\right)$$

$$= \frac{3}{2}\left[\left(1-\frac{1}{3}\right) - \left(0.5 - \frac{(0.5)^3}{3}\right)\right]$$

$$= 0.3125.$$

Q: Probability density function of a random variable X $f(x) = \begin{cases} \frac{1}{2}\sin x & 0 \le x \le \pi \\ 0 & \text{otherwise} \end{cases}$

Find mean, mode, median of the

Scanned with CamScanner

distribution and also find the probability blw 0 and $\pi/2$

Mean

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

$$\int_{-\infty}^{0} x \frac{1}{2}\sin x\,dx + \int_{0}^{\pi} x \frac{1}{2}\sin x\,dx + \int_{\pi}^{+\infty} x \frac{1}{2}\sin x\,dx = 1$$

$$0 + \int_{0}^{\pi} x \frac{1}{2}\sin x\,dx + 0 = 1$$

$$\frac{1}{2}\int_{0}^{\pi} x \sin x\,dx.$$

$$= \frac{1}{2}\left[x(-\cos x) - 1\cdot(-\sin x)\right]_0^{\pi}$$

$$= \frac{1}{2}\left[\pi(-\cos\pi) + \sin\pi) - (0 + \sin 0)\right]$$

$$= \frac{1}{2}\left[\pi(1) + 0\right]$$

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

Mode

Mode is the value of x for which $f(x)$ is maximum

$$f(x) = \frac{1}{2}\sin x$$

$$f'(x) = \frac{1}{2}\cos x$$

$f(x)$ is maximum $f'(x)$ is to be max. then $f'(x) = 0$

Scanned with CamScanner

$\Rightarrow \frac{1}{2}\cos x = 0$

$\cos x = 0$

$\cos x = \cos \pi/2$

$x = \pi/2$

$f''(x) = -\frac{1}{2}\sin x$

$= -\frac{1}{2}\sin \pi/2$

$= -\frac{1}{2} < 0$

$f(x)$ is max. at $x = \pi/2$

(iii) If M is the median of the distribut^n then,

$$\int_a^M f(x)dx = \int_M^b f(x)\,dx = 1/2$$

$$\int_0^M f(x)\,dx = 1/2$$

$$\int_0^M \frac{1}{2}\sin x\,dx = \frac{1}{2}$$

$$\frac{1}{2}\left[-\cos x\right]_0^M = \frac{1}{2}$$

$$-\cos(M) + \cos 0 = 1$$

$$-\cos M = 0$$

$$M = \pi/2.$$

Scanned with CamScanner

$$P(0 < x < \pi/2) = \int_0^{\pi/2} f(x)\,dx$$

$$= \int_0^{\pi/2} \frac{1}{2}\sin x\,dx$$

$$= \frac{1}{2}[-\cos x]_0^{\pi/2}$$

$$= \frac{1}{2}\left[-\cos\frac{\pi}{2} + \cos 0\right]$$

$$= \frac{1}{2}(0+1) = 1/2$$

$\therefore$ Mean = Mode = Median = $1/2$

Q: If x is a continuous random variable $Y = ax+b$. Prove that $E(Y) = aE(x)+b$ and $V(Y) = a^2V(x)$ where V stands for variance and a, b are constants.

Sol: Given that,

$$Y = ax+b$$

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

$$\left. E(\phi(x)) = \int_{-\infty}^{\infty} \phi(x) f(x)\,dx \right]$$

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

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

$$= \int_{-\infty}^{\infty} ax f(x)\,dx + \int_{-\infty}^{\infty} b f(x)\,dx$$

Scanned with CamScanner

$$= a\int_{-\infty}^{\infty} x f(x)\,dx + b\int_{-\infty}^{\infty} f(x)\,dx$$

$$= a\,E(x) + b \times 1$$

$$= a\,E(x) + b$$

→ let $Y = ax + b$ — ①

$$E(Y) = E(ax + b)$$

$$E(Y) = aE(x) + b \quad — ②$$

from ① − ②

$$Y - E(Y) = ax + b - aE(x) - b$$

$$Y - E(Y) = ax - aE(x)$$

S.O.B.S

$$[Y - E(Y)]^2 = a^2[x - E(x)]^2$$

Again taking E on B.S

$$E[Y - E(Y)]^2 = a^2 E[(x - E(x)]^2$$

$$V(Y) = a^2 V(x)$$

q: If x is a continuous Random variable and k is constant then P.T

i) $V(x+k) = V(x)$

ii) $V(kx) = k^2 V(x)$

Sol: WKT,

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

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

Scanned with CamScanner

PROBABILITY DISTRIBUTIONS.

→ consider a random experiment that has only two possible outcomes for example, In an experiment of rolling a die sucess corresponds to a getting a die number 1 on the face of die. Failure corresponds to getting 2,3,4,5,6, on the face of die.

→ Assume that the probability of getting success is 'p' and failing is 'q' such that $\boxed{p+q=1}$

Sometimes we need to find the probability of 'x' success among 'n' trials under the following conditions.

i) There are 'n' trials which are independent
ii) There are only '2' possible outcomes that is p & q (for each trial)
iii) The probability of sucess is same for each and every trial.

BERNOULI'S DISTRIBUTION:

A random variable X which takes two values 0 & 1 with probabilities q & p such that $P(x=0)=q$ & $P(x=1)=P$

where, $\boxed{q=1-P}$ is called Bernoulis discrete random variable and it is said to have Bernouli's distribution.

Bernouli's distribution can be written as

Scanned with CamScanner

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

BINOMIAL DISTRIBUTION

Binomial distribution was discovered by James bernouli in the year of 17th century. consider a series of 'n' independent trials in which the probability of success is P and probability of failure is q in any trial.

→ The probability of x success consequently (n-x) failures in 'n' trials.

$$= p^x q^{n-x}$$

→ But 'x' success among 'n' trials can occurs nc_x ways and the probability of each way is $p^x q^{n-x}$

→ The probability x success in 'n' trials in any order is given by $nc_x p^x q^{n-x}$

DEFINATION

A random variable X is said to follow binomial distribution, if it assume only non-negative values and its probability distribution is given by

$$P(x=x) = nc_x\, p^x q^{n-x}$$

$$= nc_x\, p^x (1-p)^{n-x}$$

where $x = 1, 2, 3 \ldots n$ and its denoted by X or $B(n, x, p)$

Scanned with CamScanner

Constriants of Binomial distribution.

The binomial probability distribution is given by $p(x) = nC_x p^x q^{n-x}$ where $x = 0, 1, 2 \cdots n$

Mean:

$$\text{Mean}(\bar{u}) = E(x) = \sum_{x=0}^{n} x\,p(x)$$

$$= \sum_{x=0}^{n} x \cdot nC_x\, p^x q^{n-x}$$

$$= 0 + nC_1 \cdot pq^{n-1} + 2\,nC_2\, p^2 q^{n-2} + \cdots + n \cdot nC_n\, p^n q^{n-n}$$

$$= n \cdot pq^{n-1} + 2\,\frac{n!}{(n-2)!2!} p^2 q^{n-2} + 3\,\frac{n!}{(n-3)!3!} p^3 q^{n-3} + \cdots np^n$$

$$= npq^{n-1} + n(n-1)p^2 q^{n-2} + n(n-1)(n-2)p^3 q^{n-3} + \cdots np^n$$

$$= np\left[q^{n-1} + (n-1)pq^{n-2} + \frac{(n-1)(n-2)p^2 q^{n-3}}{2} + \cdots p^{n-1}\right]$$

$$= np[q+p]^{n-1}$$

$$= np[1-p+p]^{n-1}$$

$$\boxed{E(x) = \mu = np}$$

Variance:

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

$$\left(E(x) = \sum_{x=0}^{n} x\,P(x)\right)$$

$$= \sum_{x=0}^{n} x^2 P(x) - (\mu)^2$$

$$= \sum_{x=0}^{n} (x^2 + x - x)\,P(x) - (\mu)^2$$

$$= \sum_{x=0}^{n} ((x^2 - x) + x)\,P(x) - \mu^2$$

$$= \sum_{x=0}^{n} [x(x-1) + x]\,P(x) - \mu^2$$

Scanned with CamScanner

$$= \sum_{x=0}^{n} x(x-1)P(x) + \sum_{x=0}^{n} xP(x) - \mu^2$$

$$= \sum_{x=0}^{n} x(x-1)P(x) + \mu - \mu^2$$

$$= \sum_{x=0}^{n} x(x-1)\, {}^nC_x p^x q^{n-x} + \mu - \mu^2$$

$$= [0+0+2\,{}^nC_2 p^2 q^{n-2} + 6\,{}^nC_3 p^3 q^{n-3} + \cdots n(n-1)\,{}^nC_n p^n q^{n-n} + \mu - \mu^2]$$

$$= \left[2\frac{n!}{(n-2)!n!}p^2q^{n-2} + 6\frac{n!}{(n-3)!3!}p^3q^{n-3} + \cdots n(n-1)p^n\right] + \mu - \mu^2$$

$$= \left[2\frac{n(n-1)(n-2)!}{(n-2)!\,2!}p^2q^{n-2} + 6\frac{n(n-1)(n-2)(n-3)!}{(n-3)!\,3!}p^3q^{n-3} + \cdots n(n-1)p^n\right] + \mu - \mu^2$$

$$= [n(n-1)p^2q^{n-2} + n(n-1)(n-2)p^3q^{n-3} + \cdots + n(n-1)p^n] + \mu - \mu^2$$

$$= n(n-1)p^2[q^{n-2} + (n-2)pq^{n-3} + \cdots + p^{n-2}] + \mu - \mu^2$$

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

$$= n(n-1)p^2[(1+p-p)^{n-2}] + \mu - \mu^2 \quad [\because q = 1-p]$$

$$= n(n-1)p^2 + \mu - \mu^2$$

$$= n^2p^2 - np^2 + \mu - \mu^2$$

$$= n^2p^2 - np^2 + np - np^2 \quad [\because \mu = np]$$

$$= np - np^2 \Rightarrow np(1-p) \quad [\because q = 1-p]$$

$$= npq$$

$$=$$

Scanned with CamScanner

Mode of binomial distribution.

Mode of Binomial distribution is a value of x at which $P(x)$ has maximum value

$$\text{Mode} = \begin{cases} \text{Integral part of } (n+1)p; \text{ if } (n+1)p \text{ is not an integer} \\ (n+1)p \text{ \& } (n+1)(p-1); \text{ if } (n+1)p \text{ is integer} \end{cases}$$

Recurrente Relation for Binomial Distribution

WKT,

$$P(x) = {}^nC_x\, p^x q^{n-x} \quad \text{— (1)}$$

$$P(x+1) = {}^nC_{x+1}\, p^{x+1} q^{n-x-1} \quad \text{— (2)}$$

eqn 2 ÷ 1

$$\frac{P(x+1)}{P(x)} = \frac{{}^nC_{x+1}\, p^{x+1} q^{n-x-1}}{{}^nC_x\, p^x q^{n-x}}$$

$$= \frac{{}^nC_{x+1}}{{}^nC_x} \cdot \frac{p^{x+1}}{p^x} \cdot \frac{q^{n-x-1}}{q^{n-x}}$$

$$= \frac{{}^nC_{x+1}}{{}^nC_x} \cdot \frac{p^x p}{p^x} \cdot \frac{q^n q^{-x} q^{-1}}{q^n q^{-x}}$$

$$= \frac{{}^nC_{(x+1)}}{{}^nC_x}\, p \cdot q^{-1}$$

$$= \frac{n!}{(n-x-1)!(x+1)!} \times \frac{(n-x)!\,x!}{n!} \times \frac{p}{q}$$

$$\frac{P(x+1)}{P(x)} = \frac{(n-x)!\,x!}{(n-x-1)!\,(x+1)!} \times \frac{p}{q}$$

Scanned with CamScanner

$$\frac{P(x+1)}{P(x)} = \frac{(n-x)(n-x-1)!\,x!}{(n-x-1)!\,(x+1)\,x!} \times \frac{p}{q}$$

$$\boxed{\frac{P(x+1)}{P(x)} = \frac{n-x}{(x+1)} \times \frac{p}{q}}$$

1) 10 coins are thrown simultaneously find probability of getting

i) Atleast 7 heads.

ii) Atmost 3 heads.- $P(X \leq 3)$

Sol: The total no. of coins $= 10 = n$

p = Probability of getting success $(H) = 1/2$

q = Probability of getting failure $(T) = 1/2$

W.K.T,

$$P(X=x) = {}^{n}C_x\, p^x q^{n-x}$$

i) Atleast 7 heads $(P(X) \geq 7)$

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

$$= P(X=7) + P(X=8) + P(X=9) + P(X=10)$$

$$= {}^{10}C_7\, p^7 q^{10-7} + {}^{10}C_8\, p^8 q^{10-8} + {}^{10}C_9\, p^9 q^{10-9} + {}^{10}C_{10}\, p^{10} q^{10-10}$$

$$= 120\left(\frac{1}{128}\right)\left(\frac{1}{8}\right) + 45\left(\frac{1}{256}\right)\left(\frac{1}{4}\right) + 10\left(\frac{1}{512}\right)\left(\frac{1}{2}\right) + 1\left(\frac{1}{1024}\right)(1)$$

$$= \frac{15}{128} + \frac{45}{1024} + \frac{10}{1024} + \frac{1}{1024}$$

$$= \frac{176}{1024} = \frac{11}{64}$$

Scanned with CamScanner

13/8/19

Q: If 10% of the rivets produced by a machine are defective, find the probability that out of 5 rivets choosen at random

(i) None will be defective

(ii) One will be defective

(iii) Atmost 2 rivets will be defective

Sol: Given,

The total no. of rivets = 5

P = Probability of defective items

$$P = \frac{10}{100} = 0.1$$

$$q = 1 - P = 1 - 0.1 = 0.9$$

W.K.T

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

(i) None will be defective

$$P(X = 0) = {}^{5}C_{0}\, p^{0} q^{5-0}$$

$$= {}^{5}C_{0}\, (0.1)^{0} (0.9)^{5}$$

$$= 1 \times 1 (0.9)^{5}$$

$$= 0.59$$

(ii) one will be defective

$$P(X = 1) = {}^{5}C_{1}\, (0.1)^{1} (0.9)^{5-1}$$

$$= 5(0.1)(0.9)^{4}$$

$$= 0.328$$

(iii) Atmost 2 will be defective

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

Scanned with CamScanner

$= {}^5C_0 (0.1)^0 (0.9)^5 + {}^5C_1 (0.1)^1 (0.9)^{5-1} + {}^5C_2 (0.1)^2 (0.9)^{5-2}$

$= 0.59 + 0.328 + 0.0729$

$= 0.99$

Q. Assume that 50% of all engineering students are good in mathematics. Determine the probabilities that among 18 engg. students.

(i) Exactly 10

(ii) Atleast 10

(iii) Atmost 8

(iv) Atleast 2 and atmost 9 are good in mathematics.

$2 \leq x \leq 9$

Sol: Given,

Total no. of students = 18

P = Probability of students who are good in mathematics

$= \frac{50}{100}$

$P = 0.5$

$q = 1 - P = 1 - 0.5 = 0.5$

(i) Exactly 10 are good in mathematics

$P(X=10) = {}^{18}C_{10} (0.5)^{10} (0.5)^{18-10}$

$= 43758 \times \frac{1}{1024} \times \frac{1}{256}$

$= 0.166$

Scanned with CamScanner

(ii) Atleast 10 are good in mathematics.

$$P(X \geq 10) = P(X=10) + P(X=11) + P(X=12) + P(X=13) + P(X=14) + P(X=15) + P(X=16) + P(X=17) + P(X=18)$$

$$= {}^{18}C_{10}(0.5)^{10}(0.5)^{18-10} + {}^{18}C_{11}(0.5)^{11}(0.5)^{18-11} + {}^{18}C_{12}(0.5)^{12}(0.5)^{18-12} + {}^{18}C_{13}(0.5)^{13}(0.5)^{18-13} + {}^{18}C_{14}(0.5)^{14}(0.5)^{18-14} + {}^{18}C_{15}(0.5)^{15}(0.5)^{18-15} + {}^{18}C_{16}(0.5)^{16}(0.5)^{18-16} + {}^{18}C_{17}(0.5)^{17}(0.5)^{18-17} + {}^{18}C_{18}(0.5)^{18}(0.5)^{18-18}$$

$$= (0.5)^{18}\left[{}^{18}C_{10} + {}^{18}C_{11} + {}^{18}C_{12} + {}^{18}C_{13} + {}^{18}C_{14} + {}^{18}C_{15} + {}^{18}C_{16} + {}^{18}C_{17} + {}^{18}C_{18}\right]$$

$$= (0.5)^{18}\left[43758 + 31824 + 18564 + 8568 + 3060 + 816 + 153 + 18 + 1\right]$$

$$= \frac{1}{262144}(106724) = 0.407$$

(iii) Atmost 8 are good in mathematics

$$P(X \leq 8) = P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5) + P(X=6) + P(X=7) + P(X=8)$$

$$= \left[{}^{18}C_0 + {}^{18}C_1 + {}^{18}C_2 + {}^{18}C_3 + {}^{18}C_4 + {}^{18}C_5 + {}^{18}C_6 + {}^{18}C_7 + {}^{18}C_8\right](0.5)^{18}$$

Scanned with CamScanner

$= (0.5)^{18}(106724)$

$= 0.407$

(iv) $P(2 \le x \le 9)$

$= \sum_{x=2}^{9} {}^{18}C_x \left(\frac{1}{2}\right)^{18}$

$= \left(\frac{1}{2}\right)^{18} \left[{}^{18}C_2 + {}^{18}C_3 + - - + {}^{18}C_9\right]$

Q: Probability of a newly generated virus attack the computer system will corrupt the 4 files out of 20 files which are opened in an hour. If 20 files are open in an hour find the probability that

Scanned with CamScanner

(i) Atleast 10 files are corrupted

(ii) Exactly 3

(iii) All files are corrupted

(iv) All files are safe

(v) More than 2 but not more than 5

Sol:- Given,

Total no. of files = 20

P = Probability of corrupted files

$= \frac{4}{20}$

$P = 0.2$

$q = 1 - P = 1 - 0.2 = 0.8$

WKT, $P(X = x) = {}^{n}C_{x}\, P^{x} q^{n-x}$

(i) Atleast 10 files are corrupted

$P(X \geq 10) = P(X=10) + P(X=11) + P(X=12) + P(X=13) + \cdots + P(X=20)$

$= {}^{20}C_{10}(0.2)^{10}(0.8)^{10} + {}^{20}C_{11}(0.2)^{11}(0.8)^{9} + {}^{20}C_{12}(0.2)^{12}(0.8)^{8} + {}^{20}C_{13}(0.2)^{13}(0.8)^{7} + {}^{20}C_{14}(0.2)^{14}(0.8)^{6} + {}^{20}C_{15}(0.2)^{15}(0.8)^{5} + {}^{20}C_{16}(0.2)^{16}(0.8)^{4} + {}^{20}C_{17}(0.2)^{17}(0.8)^{3} + {}^{20}C_{18}(0.2)^{18}(0.8)^{2} + {}^{20}C_{19}(0.2)^{19}(0.8)^{1} + {}^{20}C_{20}(0.2)^{20}(0.8)^{0}$

$P(X \geq 10) = (2.031 \times 10^{-3}) + (4.616 \times 10^{-4}) + (8.656 \times 10^{-5}) + (1.331 \times 10^{-5}) + (1.664 \times 10^{-6}) + (1.664 \times 10^{-7}) + (1.300 \times 10^{-8}) + (7.650 \times 10^{-10}) + (3.187 \times 10^{-11}) + (8.388 \times 10^{-13}) + (1.048 \times 10^{-14})$

Scanned with CamScanner

=

(ii) Exactly 3

$$P(X=3) = {}^{20}C_3 (0.2)^3 (0.8)^{17}$$

$$= 0.2053.$$

(iii) All files are corrupted

$$P(X=20) = {}^{20}C_{20} (0.2)^{20} (0.8)^0$$

$$= 1.048 \times 10^{-14}$$

(iv) All files are safe.

$$P(X=0) = {}^{20}C_0 (0.2)^0 (0.8)^{20}$$

$$= 0.0115$$

(v) More than 2 but not more than 5

$$P(2 < x \leq 5) = P(X=3) + P(X=4) + P(X=5)$$

$$= {}^{20}C_3 (0.2)^3 (0.8)^{17} + {}^{20}C_4 (0.2)^4 (0.8)^{16} + {}^{20}C_5 (0.2)^5 (0.8)^{15}$$

$$= 0.205 + 0.218 + 0.174$$

$$= 0.597.$$

Q: 20% of items from a factory are defective. Find the probability that in a sample of 5 items.

(i) None will be defective

(ii) One is defective

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

Sol :- Given,

Total no. of items = 5

Scanned with CamScanner

P = Probability of defective items.

$= \frac{20}{100}$

$P = 0.2$

$q = 1 - P = 1 - 0.2 = 0.8$

W.K.T, $P(X = x) = {}^{n}C_{x}\, p^{x} q^{n-x}$

(i) None will be defective

$P(X = 0) = {}^{5}C_{0}\,(0.2)^{0}\,(0.8)^{5-0}$

$= 0.327$

(ii) One is defective

$P(X = 1) = {}^{5}C_{1}\,(0.2)^{1}\,(0.8)^{5-1}$

$= 0.409$

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

$= {}^{5}C_{2}\,(0.2)^{2}(0.8)^{3} + {}^{5}C_{3}\,(0.2)^{3}(0.8)^{2}$

$= 0.204 + 0.051$

$= 0.255$

Q: The mean and variance of a binomial distribution are 4 and 4/3 resp. Find the probability of $P(x \geq 1)$

Sol: Given,

Mean, $\mu = 4$

$np = 4$

Variance, $V(x) = 4/3$

$npq = \frac{4}{3}$

$4(q) = 4/3$

Scanned with CamScanner

$q = 1/3$

$p = 1 - q = 1 - \frac{1}{3} = \frac{2}{3} \Rightarrow p = \frac{2}{3}$

$np = 4$

$n\left(\frac{2}{3}\right) = 4$

$n = \frac{12}{2}$

$n = 6$

$P(X \geq 1) = P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5) + P(X=6)$

$= {}^6C_1 \left(\frac{2}{3}\right)^1 \left(\frac{1}{3}\right)^5 + {}^6C_2 \left(\frac{2}{3}\right)^2 \left(\frac{1}{3}\right)^4 + {}^6C_3 \left(\frac{2}{3}\right)^3 \left(\frac{1}{3}\right)^3 + {}^6C_4 \left(\frac{2}{3}\right)^4 \left(\frac{1}{3}\right)^2 + {}^6C_5 \left(\frac{2}{3}\right)^5 \left(\frac{1}{3}\right)^1 + {}^6C_6 \left(\frac{2}{3}\right)^6 \left(\frac{1}{3}\right)^0$

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

$= 1 - {}^6C_0 \left(\frac{2}{3}\right)^0 \left(\frac{1}{3}\right)^6$

$= 1 - \frac{1}{729}$

$= 0.998$

Q: The mean and variance of binomial variable x with parameters n, p are 16 and 8. Find (i) $P(X \geq 1)$ (ii) $P(X > 2)$

Sol:- Given,

Mean, $\mu = 16$

$np = 16$

Variance, $v(x) = 8$

$npq = 8$

Scanned with CamScanner

$16q = 8$

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

$P = 1 - q = 1 - \frac{1}{2} = \frac{1}{2} \Rightarrow P = \frac{1}{2}$

$np = 16$

$\frac{1}{2}n = 16 \Rightarrow n = 32$

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

$= 1 - {}^{32}C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^{32}$

$= 1 - (2.328 \times 10^{-10})$

$= 0.9999999$

(ii) $P(X > 2) = 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$

$= 1 - \left[\left({}^{32}C_0 + {}^{32}C_1 + {}^{32}C_2\right)\left(\frac{1}{2}\right)^{32}\right]$

$= 1 - [(2.328 \times 10^{-10})(529)]$

$= 1 - (1.231 \times 10^{-7})$

$= 0.9999998769$

Q: In a 8 throws of a die, 5 (or) 6 is considered as success. Find the mean no. of success and standard deviation.

Sol:- Given,

Total no. of throws = 8

P = Probability of getting 5 (or) 6 when a die is thrown.

$= \frac{2}{6}$

Scanned with CamScanner

$$P = \frac{1}{3}$$

$$q = 1 - \frac{1}{3} = \frac{2}{3}$$

W.K.T,

Mean, $\mu = np$

$$= 8\left(\frac{1}{3}\right)$$

$$\boxed{\mu = \frac{8}{3}}$$

Variance, $V(x) = npq$

$$= 8 \times \frac{1}{3} \times \frac{2}{3}$$

$$V(x) = \frac{16}{9}$$

standard derivation, $\sigma = \sqrt{npq}$

$$= \sqrt{\frac{16}{9}}$$

$$\boxed{\sigma = \frac{4}{3}}$$

16|8|19

Q: In a binomial distribution consisting of 5 Independent dies probability of 1 and 2 success are 0.4096 and 0.2048 respectively. Find P of distribution.

Sol: Given,

No. of trials $(n) = 5$

$P(X = 1) = 0.4096$

$P(X = 2) = 0.2048$

Scanned with CamScanner

$$\frac{P(X=1)}{P(X=2)} = \frac{0.4096}{0.2048}$$

$$\left[P(X=x) = {}^{n}C_x\, p^x q^{n-x} = {}^{n}C_x\, (1-p)^{n-x}\, p^x\right]$$

$$\frac{{}^{5}C_1\, p^1 (1-p)^{5-1}}{{}^{5}C_2\, p^2 (1-p)^{5-2}} = \frac{0.4096}{0.2048}$$

$$\frac{5p(1-p)^4}{10p^2(1-p)^3} = 2$$

$$\frac{5(1-p)}{10p} = 2$$

$$5(1-p) = 20p$$

$$5 - 5p = 20p$$

$$\boxed{p = 0.2}$$

Q: The probability of a man hitting a target is 1/3

i) If he fires 5 times. what is the probability of hitting his target atleast twice

ii) How many he fire so that the prob. of his hitting the target atleast once is more than 90%.

Sol: Given,

P = Probability of man hitting a target = 1/3

$$q = 1 - p = 1 - \frac{1}{3} = \frac{2}{3}$$

Scanned with CamScanner

(i) $n = 5$

$$P(X \geq 2) = 1 - P(X < 2)$$

$$= 1 - [P(X=0) + P(X=1)]$$

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

$$= 1 - \left[\left(\frac{2}{3}\right)^5 + \frac{5}{3}\left(\frac{2}{3}\right)^4\right]$$

$$= 1 - \left[\frac{32}{243} + \frac{80}{243}\right]$$

$$= 1 - \frac{112}{243} = \frac{131}{243}$$

$$= 0.539$$

ii) $P(X \geq 1) = 1 - P(X=0) > 90\%$

$$= 1 - P(X=0) > 90/100$$

$$= 1 - {}^nC_0 (1/3)^0 \left(\frac{2}{3}\right)^{n-0} > 90/100$$

$$= 1 - (2/3)^n > 0.9$$

$\therefore$ This is satisfied for $n = 6$

Q: Find the maximum n such that the probability of getting no head in a tossing a fair coin n times is greater than 0.1

Sol: P = Probability of getting head

$P = 1/2$

q = probability of not getting head

$q = 1 - P = 1/2$

Scanned with CamScanner

$P(x=0) > 0.1$

$n_{C_0} (1/2)^0 (1/2)^{n-0} > 0.1$

$(1/2)^n > 0.1$

For $n=1 \Rightarrow (1/2)^1 = 0.5 > 0.1$

$n=2 \Rightarrow (1/2)^2 = 0.25 > 0.1$

$n=3 \Rightarrow (1/3)^3 = 0.125 > 0.1$

∴ The require maximum value is $n=3$.

Binomial Distribution Frequency.

If n independent trials constitute 1 experiment and this experiment is repeated an N times then the frequency of x success is $n=N$.

$n_{C_x} p^x q^{n-x}$

Since the probabilities are 0, 1, 2 ---- r ---- n success in n trials are given by the terms of binomial expansion of $(q+p)$

∴ In N sets of n trials, the theoritical frequencies of 0, 1, 2 --- r --- n success will be given by the terms of expansion of the $N(q+p)^n$. The possible no. of success and their frequencies are called as binomial frequency distribution.

Q: Fit a binomial distribution for the following frequency distribution

x	0	1	2	3	4	5	6
f	13	25	52	58	32	16	4

Scanned with CamScanner

Sol: Given that

No. of trials$(n) = 6$

Frequency, $\Sigma f_i = 200 = N$

Mean $np = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{\Sigma f_i x_i}{N}$

$$np = \frac{0+25\times1+2\times52+3\times58+4\times32+15\times16+6\times4}{13+25+52+58+32+16+4}$$

$$np = \frac{0+25+104+174+128+80+24}{200}$$

$$np = \frac{530}{200}$$

$$np = 2.65$$

$$6p = 2.65$$

$$p = 0.445$$

$q = 1-p = 1-0.445 = 0.555$

Expected frequency distribution is as follows

x	f_i	$P(x) = nCx\, p^x q^{n-x}$	Expected frequency $E(x) = N \cdot P(x) = N\, nCx\, p^x q^{n-x}$
0	13	0.0305	6.1
1	25	0.1444	28.88
2	52	0.2848	56.96
3	58	0.2996	59.92
4	32	0.1772	35.44
5	16	0.0559	11.18
6	4	0.0073	1.46

Scanned with CamScanner

Expected frequency distribution.

x	0	1	2	3	4	5	6
f	13	25	52	58	32	16	4
E(x)	6	28	56	60	36	12	1

It may be noted that expected frequencies are closed to observed frequencies

Q: 4 coins are tossed 160 times the no. of times x heads occurs is given below

x	0	1	2	3	4
No. of times	8	34	69	43	6

Fit a Binomial distribution to this data on the hypothesis that coins are unbiased.

Scanned with CamScanner

20/8/19

POISSION DISTRIBUTION FUNCTION

Poission Distribution is a limiting case of binomial distribution under the following conditions.

→ n is very large (i.e $n \to \infty$)
→ p is very small (i.e $p \to 0$)
→ $np = \lambda$ (λ is finite)

Derivation of Poission Distribution

Sol:

In the Binomial distribution probability $P(x)$ of x success in a series of independent trials is given by

$$P(X=x) = {}^nC_x\, p^x q^{n-x}$$

$$= {}^nC_x\, p^x (1-p)^{n-x}$$

$$= \frac{n!}{(n-x)!\,x!}\, p^x (1-p)^{n-x}$$

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

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

WKT,

$$np = \lambda \Rightarrow n = \lambda/p$$

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

Scanned with CamScanner

$$= \frac{\lambda(\lambda-P)(\lambda-2P)\cdots[\lambda-P(x-1)]\,P^x(1-P)^{\lambda/P}}{x!\,P^x\,(1-P)^x}$$

$$= \frac{\lambda(\lambda-P)(\lambda-2P)\cdots[\lambda-P(x-1)]\,(1-P)^{\lambda/P}}{x!\,(1-P)^x}$$

P is very small i.e $P \to 0$

$$= \frac{\lambda(\lambda)(\lambda)\cdots\lambda\ (x \text{ times})}{x!\,(1-P)^x}(1-P)^n$$

$$= \frac{\lambda^x}{x!}\frac{(1-P)^n}{(1-P)^x}$$

$$= \frac{\lambda^x}{x!} \times \lim_{n\to\infty} \frac{(1-\lambda/n)^n}{(1-\lambda/n)^x}$$

$$= \frac{\lambda^x}{x!} \times \lim_{n\to\infty} \frac{(1-\lambda/n)^n}{(1-0)^x}$$

$$= \frac{\lambda^x}{x!} \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n$$

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

If the probability function of P.D where $x = 0, 1, 2, \cdots\cdots$

DEFINATION.

A random variable X is said to follow P.D if it assumes only non -ve values and its probability density function is given by

Scanned with CamScanner

$$P(X=x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!}, & x = 0,1,2,\ldots \quad \lambda > 0 \\ 0 & \text{otherwise.} \end{cases}$$

CONSTRAINTS OF P.D.F

1) Mean of P.D

W.K.T,

$$\text{Mean}(\mu) = \Sigma x P(x)$$

$$= \sum_{x=0}^{n} x \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{n} x \frac{e^{-\lambda}\lambda^x}{x(x-1)!}$$

$$= e^{-\lambda} \sum_{x=0}^{n} \frac{\lambda^x}{(x-1)!}$$

$$= e^{-\lambda} \sum_{x=1}^{n} \frac{\lambda^x}{(x-1)!}$$

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

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

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

$$\left[\because e^{tx} = \left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots\right)\right]$$

$$e^{-x} = \left(1 - \frac{x}{1!} + \frac{x^2}{2!} - \ldots\right)$$

$$\mu = \lambda$$

$$\Rightarrow \bar{\mu} = E(X) = \bar{x} = \lambda$$

Scanned with CamScanner

$\Rightarrow \bar{x} = \mu = d$

$\boxed{np = d}$

Thus the parameter d is arithmatic mean of P.D

Variance of P.D

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

$$= E(x^2) - (\bar{x})^2$$

$$= \Sigma x^2 P(x) - (\bar{x})^2$$

$$= \Sigma (x^2 + x - x) P(x) - (\bar{x})^2$$

$$= \Sigma ((x^2 - x) + x) P(x) - (\bar{x})^2$$

$$= \Sigma [x(x-1) + x] P(x) - (\bar{x})^2$$

$$= \Sigma x(x-1) P(x) + \Sigma x P(x) - \bar{x}^2$$

$$= \Sigma x(x-1) P(x) + d - d^2$$

$$= \sum_{x=0}^{n} x(x-1) \frac{e^{-d} d^x}{x!} + d - d^2$$

$$= 0 + 0 + \sum_{x=0}^{n} x(x-1) \frac{e^{-d} d^x}{x(x-1)(x-2)!} + d - d^2$$

$$= \sum_{x=0}^{n} \frac{e^{-d} d^x}{(x-2)!} + d - d^2$$

$$= \sum_{x=2}^{n} \frac{e^{-d} d^x}{(x-2)!} + d - d^2$$

$$= e^{-d} \sum_{x=2}^{n} \frac{d^x}{(x-2)!} + d - d^2$$

Scanned with CamScanner

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

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

$$= e^{-\lambda}\lambda^2(e^{\lambda}) + \lambda - \lambda^2 = \lambda^2 + \lambda - \lambda^2$$

$$V(x) = \lambda$$

∴ variance of P.D is λ

→ standard deviation of P.D $= \sqrt{\lambda}$

MODE OF THE P.D

Mode is the value of x in which the probability P(x) is maximum

we have,

$$P(x) \geq P(x+1) \text{ (or) } P(x) \geq P(x-1)$$

$$\frac{e^{-\lambda}\lambda^x}{x!} \geq \frac{e^{-\lambda}\lambda^{(x+1)}}{(x+1)!} \text{ (or) } \frac{e^{-\lambda}\lambda^x}{x!} \geq \frac{e^{-\lambda}\lambda^{(x-1)}}{(x-1)!}$$

$$\frac{e^{-\lambda}\lambda^x}{x!} \geq \frac{e^{-\lambda}\lambda^x \cdot \lambda}{(x+1)x!} \text{ (or) } \frac{e^{-\lambda}\lambda^x}{x(x-1)!} \geq \frac{e^{-\lambda}\cdot e\lambda^x \cdot \lambda^{-1}}{(x-1)!}$$

$$1 \geq \frac{\lambda}{x+1} \quad \text{(or)} \quad \frac{1}{x} \geq \frac{1}{\lambda}$$

$$x+1 \geq \lambda$$

$$x \geq \lambda - 1 \quad \text{(or)} \quad x \leq \lambda$$

Hence mode of P.D lies b/w $(\lambda-1)$ & λ

Scanned with CamScanner

Recurrance relation for P.D

W.K.T,

$$P(x) = \frac{e^{-d} d^x}{x!} \quad \text{---} \quad (1)$$

$$P(x+1) = \frac{e^{-d} d^{(x+1)}}{(x+1)!} \quad \text{---} \quad (2)$$

$$\frac{P(x+1)}{P(x)} = \frac{e^{-d} d^{x+1}}{(x+1)!} \times \frac{x!}{e^{-d} d^x}$$

$$= \frac{e^{-d} d^x \cdot d}{(x+1)x!} \times \frac{(x+1)\, x!}{e^{-d} d^x}$$

$$\frac{P(x+1)}{P(x)} = \frac{d}{(x+1)}$$

$$P(x+1) = \frac{d \cdot P(x)}{(x+1)} \quad (\text{or})$$

$$P(x) = \frac{x+1}{d} P(x+1)$$

Q: A manufacture of cotton pins knowns that 5% of its product its defective. Pins are sold in boxes of 100. what is the probability more than 3 pins is defective

Sol: Given,

P = Probability of defective = $5\% = \frac{5}{100}$

$P = 0.05$

$n = 100$

Scanned with CamScanner

W.K.T, mean(μ) $= np = \lambda$

$\lambda = 100 \times 0.05$

$\lambda = 5$

W.K.T, $P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$

$= \frac{e^{-5}5^x}{x!}$

$P(X>3) = 1-[P(X=0)+P(X=1)+P(X=2)+P(X=3)]$

$= 1-\left[\frac{e^{-5}5^0}{0!}+\frac{e^{-5}5^1}{1!}+\frac{e^{-5}5^2}{2!}+\frac{e^{-5}5^3}{3!}\right]$

$= 1-e^{-5}\left[1+\frac{5}{1!}+\frac{5^2}{2!}+\frac{5^3}{3!}\right]$

$= 1-e^{-5}\left[1+5+\frac{25}{2}+\frac{125}{6}\right]$

$= 1-e^{-5}\left[\frac{118}{3}\right]$

$P(X=x) = 0.734$

23/9/19

Q: 2% of the items of a factory are defective. The items are packed in boxes. What is the probability that there will be

(i) 2 defective items

(ii) atleast 3 defective items in a box of 100 items

Sol : Given,

$n = 100$

p = probability of defective items

Scanned with CamScanner

$= \frac{2}{100}$

$p = 0.02$

$q = 1 - p = 1 - 0.02 = 0.98.$

W.K.T

$\text{Mean}(\mu) = np = \lambda$

$= 100 \times 0.02$

$\lambda = 2.$

(i) we have $P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$

$= \frac{e^{-2} 2^x}{x!}$

Given 2 are defective items

$x = 2.$

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

$= 0.2706.$

(ii) Atleast 3 defective items.

$P(x \geq 3) = 1 - [P(X=0) + P(X=1) + P(X=2)]$

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

$= 1 - e^{-2}[1 + 2 + 2]$

$= 1 - 5e^{-2}$

$= 0.3233$

Scanned with CamScanner

Q: Average no. of accidents on an day on a national highway is 1.8. Determine the probability that the no. of acidents are (i) Atleast 1

(ii) Atmost 1

Given,

Average no. of accidents $= 1.8$

Mean $(np) = \lambda = 1.8$

(i) Atleast 1

WKT, $P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$

$1 - P(X=0)$

$$P(x \geq 1) = 1 - \frac{e^{-1.8}\,1.8^0}{0!}$$

$$= 1 - e^{-1.8}\,1^{1.8}$$

$$= 0.834$$

(ii) Atmost 1

$$P(X \leq 1) \Rightarrow P(X=0) + P(X=1)$$

$$= \frac{e^{-1.8}\,1.8^0}{0!} + \frac{e^{-1.8}\,1.8^1}{1!}$$

$$(= 0.834 + 0.297$$

$$= 1.131)$$

$$= e^{-1.8}(1 + 1.8)$$

$$= 0.4628$$

Q: If a random variable has a P.D

$P(1) = P(2)$

Scanned with CamScanner

find

(i) Mean (ii) $P(4)$ (iii) $P(1<x<4)$ (iv) $P(x \geq 1)$

Sol:

given;

$P(1) = P(2)$

we have, $P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$

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

$$1 = \frac{\lambda}{2} \Rightarrow \lambda = 2.$$

(i) Mean.

Mean $(np) = \lambda$

$\lambda = 2.$

(ii) $P(4)$

$$P(X=4) = \frac{e^{-2}2^4}{4!}$$

$$= 0.090.$$

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

$$\Rightarrow P(x=2) + P(x=3)$$

$$= \frac{e^{-2}2^2}{2!} + \frac{e^{-2}2^3}{3!} \Rightarrow e^{-2}\left[\frac{4}{2} + \frac{8}{6}\right]$$

$$= e^{-2}\left[2 + \frac{4}{3}\right]$$

$$= 0.4511$$

Scanned with CamScanner

(iv) $P(X \geq 1)$

$= 1 - P(X=0)$

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

$= 0.864$

Q: If a P.D is such that $P(X=1)\frac{3}{2} = P(X=3)$. Find

(i) $P(X \geq 1)$ (ii) $P(X \leq 3)$ (iii) $P(2 \leq x \leq 5)$

Given,

$P(X=1)\frac{3}{2} = P(X=3)$

$\left(\frac{e^{-\lambda}\lambda^1}{1!}\right)3 = 2\left(\frac{e^{-\lambda}\lambda^3}{3!}\right)$

$(e^{-\lambda}\lambda)\,3 = \frac{e^{-\lambda}\lambda^3}{3}$

$3 = \frac{\lambda^2}{3} \Rightarrow \lambda^2 = 9$

$\lambda = \pm 3$

$\therefore \boxed{\lambda = 3}$

(i) $P(X \geq 1)$

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

$= 1 - \frac{e^{-3} 3^0}{0!}$

$= 1 - e^{-3}$

$= 0.950$

(ii) $P(X \leq 3)$

$= P(X=0) + P(X=1) + P(X=2) + P(X=3)$

Scanned with CamScanner

$$= \frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!}$$

$$= e^{-3}\left[1 + 3 + \frac{9}{2} + \frac{27}{6}\right]$$

$$= 0.64723$$

(iii) $P(2 \le x \le 5)$

$$P(x=2) + P(x=3) + P(x=4) + P(x=5)$$

$$= \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!} + \frac{e^{-3}3^4}{4!} + \frac{e^{-3}3^5}{5!}$$

$$= e^{-3}\left[\frac{9}{2} + \frac{27}{6} + \frac{81}{24} + \frac{243}{120}\right]$$

$$= 0.7169$$

Q: If the variance of P.D is 3, find the probability that

(i) $x = 0$

(ii) $0 < x \le 3$

(iii) $1 \le x < 4$

Given,

$$V(x) = 3$$

$$\therefore \lambda = 3$$

(i) $P(x=0)$ we have, $\frac{e^{-\lambda}\lambda^x}{x!}$

$$\frac{e^{-3}3^0}{0!} \Rightarrow 0.0497$$

(ii) $0 < x \le 3$

$$P(x=1) + P(x=2) + P(x=3)$$

Scanned with CamScanner

$$\frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!}$$

$$= e^{-3}\left[3 + \frac{9}{2} + \frac{27}{6}\right]$$

$$= 0.5974$$

(iii) $1 \le x < 4$

$$P(x=1) + P(x=2) + P(x=3)$$

$$= \frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!}$$

$$= e^{-3}\left[3 + \frac{9}{2} + \frac{27}{6}\right]$$

$$= 0.5974$$

Q: If two cards are drawn from a pack of 52 cards which are diamonds, using P.D find the probability of getting 2 diamonds atleast 3 times in 51 consecutive trials of 2 cards drawing each trial.

Sol:

P = Probability of getting 2 diamonds from a pack of 52 cards

$$= \frac{13C_2}{52C_2} = \frac{3}{51}$$

$$= \frac{1}{17}$$

∴ No. of trials $(n) = 51$

Mean $(np) = d$

Scanned with CamScanner

$$\Rightarrow 51 \times \frac{1}{17} = \lambda$$

$$\lambda = 3.$$

From P.D

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

Atleast 3 times

$$P(X \geq 3) = 1 - [P(X=0) + P(X=1) + P(X=2)]$$

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

$$= 1 - e^{-3}\left[1 + 3 + \frac{9}{2}\right]$$

$$= 1 - e^{-3}\left[4 + \frac{9}{2}\right]$$

$$= 0.5768$$

Q: Suppose 2% of people on the average are left handed.

(i) Find the probability of finding 3 or more left handed.

(ii) The prob. of finding none are left handed.

Sol:

Given,

average people(λ) = 2% = $\frac{2}{100}$

$$\lambda = 0.02$$

(i) $P(X \geq 3) = 1 - [P(X=0) + P(X=1) + P(X=2)]$

$(= 1 - [e^{-0.02}(0.02)^0)$ $+ P(X=3)]$

Scanned with CamScanner

$$= 1 - \left[\frac{e^{-0.02}(0.02)^0}{0!} + \frac{e^{-0.02}(0.02)^1}{1!} + \frac{e^{-0.02}(0.02)^2}{2!} \right]$$

$$= 1.313 \times 10^{-6}.$$

(iii) $P(0 < x \leq 1) = P(X=1)$

$$= \frac{e^{-0.02}(0.02)^1}{1!}$$

$$= 0.0196$$

Q: If a bank recieved on the avg. six bad cheques per day, Find the prob. that it will receive 4 bad cheques on any given day.

Sol:

Average bad cheques $(d) = 6$.

W.K.T $P(X=x) = \frac{e^{-d} d^x}{x!}$

$$P(X=4) = \frac{e^{-6} e^4}{4!}$$

$$= 0.1338$$

Q: A person receives an avg. of 4 telephone calls per day. Assuming P.D what is the prob. that on a given day he or (may) will rcieve

(i) no telephone calls

(ii) exactly one call

(iii) Atleast 2 telephone calls.

Scanned with CamScanner

Sol : Given ,

$\lambda = 4$

W.K.T $P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$

(i) $P(X=0)$

$$\frac{e^{-4}4^0}{0!}$$

$$= 0.0183$$

(ii) $P(=1)$

$$\frac{e^{-4}4^1}{1!}$$

$$= 0.073$$

(iii) $P(X \geq 2)$

$$= 1 - [P(X=0) + P(X=1)]$$

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

$$= 1 - e^{-4}[1+4]$$

$$= 0.908$$

Q: If the probability that an individual suffers a bad reaction from a certain injection is 0.001. Determine the prob. that out of 2000 individual

(i) exactly 3

(ii) More than 2 individual

(iii) None

(iv) More than 1 individual suffers

Scanned with CamScanner

Q: It has been found that 2% of the tools produced by a certain machine are defective. What is the probability that in a shipment of 400 such tools

a) 3% or more

b) 2% or less will prove defective.

Q: P.T the 3 successive values of a poission variate can not have equal Probability of success.

27/8/19

Sol:- let $x-1, x, x+1$ be 3 successive poission variants

we have to show that,

$$P(x-1) \neq P(x) \neq P(x+1)$$

Now, let us assume that,

$$P(x-1) = P(x) = P(x+1)$$

1) $P(x) = P(x-1)$

$$\frac{e^{-d} d^x}{x!} = \frac{e^{-d} d^{(x-1)}}{(x-1)!}$$

$$\frac{d^x}{x(x-1)!} = \frac{d^x}{d(x-1)!}$$

$$\boxed{d = x} \rightarrow (1)$$

2) $P(x+1) = P(x)$

$$\frac{e^{-d} d^{x+1}}{(x+1)!} = \frac{e^{-d} d^x}{x!}$$

Scanned with CamScanner

$$\frac{e^{-\lambda}\lambda^{x+1}}{\lambda(x+1)x!} = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$\lambda = x+1 \longrightarrow (2)$$

From (1) & (2)

$$x = x+1 = \lambda$$

∴ Our assumption is wrong since x and $x+1$ are 2 distinct successive values and λ is unique.

∴ 3 successive values of poission variate cannot have equal probability of success.

Q: Using recurrance relation find the probability when $x = 0, 1, 2, 3, 4$ of the mean of poission distribution is 3.

Sol: W.K.T,

recurrance relation in P.D

$$P(x+1) = \frac{\lambda}{x+1} P(x)$$

given that mean $= \lambda = 3$

$$x=0,\ P(0+1) = \frac{\lambda}{0+1} P(x=0)$$

$$= 3\frac{e^{-3}3^{0}}{0!}$$

$$P(1) = 3e^{-3}$$

$$x=1,\ P(1+1) = \frac{\lambda}{1+1} P(x=1)$$

$$= \frac{3}{2}\frac{e^{-3}3^{1}}{1!} = \frac{9}{2}e^{-3}$$

Scanned with CamScanner

$$x = 2,\ P(2+1) = \frac{3}{2+1} P(X=2)$$

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

$$P(3) = \frac{e^{-3}9}{2}$$

$$x = 3,\ P(3+1) = \frac{3}{4} \frac{e^{-3}3^3}{3!}$$

$$= \frac{3}{4} \frac{e^{-3}27}{3\times 2}$$

$$P(4) = \frac{27}{8} e^{-3}$$

$$x = 4,\ P(4+1) = \frac{3}{5} \frac{e^{-3}3^4}{4!}$$

$$= \frac{3}{5} \frac{e^{-3}81}{24}$$

$$P(5) = \frac{81}{40} e^{-3}$$

Q : Fit a P.D to the following data

x	0	1	2	3	4	5	Total
$f(x)$	142	156	69	27	5	1	400

Sol:

Given P.D

x	0	1	2	3	4	5	Total
$f(x)$	142	156	69	27	5	1	400

$$\text{Mean} = np = \frac{\Sigma x_i f_i}{N}$$

$N = \Sigma f_i$

$$= \frac{156 + 138 + 81 + 20 + 5}{400}$$

Scanned with CamScanner

$np = 1$

$\lambda = 1$

$$\rightarrow P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$P(X=0) = \frac{e^{-1}1^0}{0!} = \frac{1}{e} = 0.3678$$

$$P(X=1) = \frac{e^{-1}1^1}{1!} = \frac{1}{e} = 0.3678$$

$$P(X=2) = \frac{e^{-1}1^2}{2!} = \frac{1}{2e} = 0.183$$

$$P(X=3) = \frac{e^{-1}1^3}{3!} = \frac{1}{6e} = 0.061$$

$$P(X=4) = \frac{e^{-1}1^4}{4!} = \frac{1}{24e} = 0.015$$

$$P(X=5) = \frac{e^{-1}1^5}{5!} = \frac{1}{120e} = 3.065 \times 10^{-3}$$

Expected F.D is as follows.

x	f_i	$P(X) = \frac{e^{-\lambda}\lambda^x}{x!}$	$N \cdot \frac{e^{-\lambda}\lambda^x}{x!}$ $N \cdot P(X=x)$
0	142	0.3678	$147.12 \simeq 147$
1	156	0.3678	$147.12 \simeq 147$
2	69	0.183	$73.56 \simeq 74$
3	27	0.061	$24.4 \simeq 25$
4	5	0.015	$6 = 6$
5	1	3.065×10^{-3}	$1.226 \simeq 1.0$
			400

Scanned with CamScanner

NORMAL DISTRIBUTION.

Random variables.

- Discrete R.V
  - B.D
  - P.D
- continuous R.V
  - Normal.D

Normal Distribution is a continuous Dist. in which the variant can take all the value with a given range is called as Normal Distribution.

Ex:

1. Height of students in class
2. Speed of vehicles.

→ It is another limiting case from the B.D

→ For large value of n when neither P nor q is very small. The N.D is divided from the B.D by increasing the no. of trials indefinitely.

DEFINATION.

Random variable X is said to have a N.D if its Probability density function is given by

$$f(x, a, b) = \frac{1}{b\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2}$$

where,

a, b are 2 parameters such that

Scanned with CamScanner

1. $-\infty < x < \infty$

2. $-\infty < a < \infty$

3. $b > 0$

Mean of N.D

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

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

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

let, $\frac{x-a}{b} = z \Rightarrow x = a + bz$

$$1 = 0 + b\frac{dz}{dx}$$

$$dx = b\,dz$$

$x = -\infty \Rightarrow z = -\infty$

$x = \infty \Rightarrow z = \infty$

$$= \frac{1}{b\sqrt{2\pi}} \int_{-\infty}^{\infty} (a+bz)\, e^{-\frac{1}{2}(z)^2} b\,dz$$

$$= \frac{1}{b\sqrt{2\pi}} \left[\int_{-\infty}^{\infty} \underbrace{a e^{-\frac{z^2}{2}}}_{\text{even funct}} dz + \int_{-\infty}^{\infty} \underbrace{b z e^{-\frac{z^2}{2}}}_{\text{odd func.}} dz \right]$$

$$= \frac{1}{\sqrt{2\pi}} \left[2\int_{0}^{\infty} a e^{-z^2/2}\, dz + 0 \right]$$

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

let, $\frac{z^2}{2} = t \Rightarrow z^2 = 2t$

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

Scanned with CamScanner

$z = \sqrt{2t}$

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

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

$z = 0 \to t = 0$

$z = \infty \to t = \infty$

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

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

$= \frac{a}{\sqrt{\pi}} \int_0^\infty e^{-t}\ t^{1/2-1}\ dt \quad \left[\because \int_0^\infty e^{-x} x^{n-1}\ dx \right.$

$\left. \sqrt{n+1} = n\sqrt{n} \right.$

$= \frac{a}{\sqrt{\pi}}\ \sqrt{1/2}$

$\left. \sqrt{1/2} = \sqrt{\pi} \right]$

$= \frac{a}{\sqrt{\pi}}\ \sqrt{\pi}$

$= a$

30/8/19

The variance of normal distribution

We know that

$\text{variance}(\sigma^2) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\ dx$

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

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

$[\mu = a]$

Scanned with CamScanner

let $\frac{x-\mu}{b} = z \Rightarrow x = \mu + bz$

$\Rightarrow 1 = 0 + b\frac{dz}{dx}$

$dx = b\,dz$

$x = -\infty \Rightarrow z = -\infty$

$x = \infty \Rightarrow z = \infty$

$$= \frac{1}{b\sqrt{2\pi}} \int_{-\infty}^{\infty} (bz)^2 e^{-\frac{1}{2}(z)^2} b\,dz$$

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

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

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

[$\int_{-a}^{a} f(x)\,dx = 0$ if 'f' is odd function

$\int_{-a}^{a} f(x) = 2\int_{0}^{a} f(x)\,dx$ if 'f' is even function]

Let $\frac{z^2}{2} = t \Rightarrow z^2 = 2t \Rightarrow z = \sqrt{2t} \Rightarrow dx = \frac{1}{z}dt$

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

$z = 0 \Rightarrow t = 0$

$z = \infty \Rightarrow t = \infty$

$$= \frac{2b^2}{\sqrt{2\pi}} \int_{0}^{\infty} 2te^{-t} \frac{dt}{\sqrt{2t}}$$

Scanned with CamScanner

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

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

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

[$\sqrt{n+1} = n\sqrt{n}$ 'n' Integer

$\sqrt{n+1} = n!$ n is +ve integer

$\sqrt{1/2} = \sqrt{\pi}$]

$$= \frac{2b^2}{\sqrt{\pi}} \int_0^{\infty} e^{-t} t^{3/2-1} \, dt$$

$$= \frac{2b^2}{\sqrt{\pi}} \sqrt{3/2}$$

$$= \frac{2b^2}{\sqrt{\pi}} \sqrt{\frac{1}{2}+1}$$

$$= \frac{2b^2}{\sqrt{\pi}} \frac{1}{2} \sqrt{1/2} = \frac{2b^2}{\sqrt{\pi}} \times \frac{1}{2}\sqrt{\pi}$$

$$\sigma^2 = b^2$$

$$\boxed{\sigma = b}$$

NOTE:

→ The standard deviation $\sigma = b$

→ The N.D with mean(μ) $\mu = a$ and the variance $\sigma^2 = b^2$ becomes

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

Scanned with CamScanner

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

Mode of Normal Distribution.

W.K.T,

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

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

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

$$= \frac{-(x-\mu)}{\sigma^3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$f(x)$ to be maximum $\Rightarrow f'(x) = 0$

$$\frac{-(x-\mu)}{\sigma^3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = 0$$

$$-(x-\mu)\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = 0$$

$$\left[\because ab = 0 \Rightarrow a = 0 \text{ (or) } b = 0\right]$$

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

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

Scanned with CamScanner

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

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

At $x=\mu$

$$f''(x) = \frac{-1}{\sigma^2\sqrt{2\pi}}\left[\frac{1}{\sigma}e^{0}(1-0)^2\right]$$

$$= -\frac{1}{\sigma^3\sqrt{2\pi}} < 0$$

$f''(x) < 0$

$\Rightarrow f(x)$ is maximum at $x=\mu$

Median of Normal Distribution.

We know that

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

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

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

consider

$$\int_{-\infty}^{\mu} f(x)\,dx + \int_{\mu}^{M} f(x)\,dx = 1/2 \longrightarrow (1)$$

Scanned with CamScanner

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

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

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

$dx = \sigma\, dz$

$x = -\infty, z = -\infty$

$x = M, z = 0$

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

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

$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{1}{2}(z)^2} \cdot \sigma\, dz$

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

let $\frac{z^2}{2} = t \Rightarrow z^2 = 2t \Rightarrow 2z\,dz = 2\,dt$

$dz = \frac{1}{z} dt$

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

$z = 0, \Rightarrow t = 0$

$z = \infty \Rightarrow t = \infty$

$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-t} \frac{dt}{\sqrt{2t}}$

Scanned with CamScanner

$$= \frac{1}{\sqrt{2\pi}} \times \frac{1}{\sqrt{\frac{1}{2}}} \int_{-\infty}^{0} e^{-t}\, dt$$

$$= \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{0})$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-t} t^{\frac{1}{2}-1}\, dt$$

$$= \frac{1}{2\sqrt{\pi}} \left|\frac{1}{2}\right.$$

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

$$= \frac{1}{2} \longrightarrow (2)$$

Sub eqn (2) in (1)

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

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

By comparing $\int_{+\infty}^{M} f(x)\, dx = 1/2$ and $\int_{-\infty}^{\mu} f(x)\, dx = 1/2$

$\therefore M = \mu$

NOTE:

→ In N.D, Mean = Mode = Median = μ

Mean Deviation from the Mean of N.D

Mean Deviation = $\int_{-\infty}^{\infty} |x - \mu|\, f(x)\, dx$

Scanned with CamScanner

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

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

(Characteristics of N.D 6/9/19

→ The graph of the N.D $y = f(x)$ in XY-plane is known as the normal curve)

$$x = \mu + \sigma z \qquad x = -\infty \Rightarrow z = -\infty$$
$$dx = \sigma dx \qquad x = \infty \Rightarrow z = \infty$$

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

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

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

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

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

$$= \frac{2\sigma}{\sqrt{2\pi}} \int_{0}^{\infty} \sqrt{2t} e^{-t} \frac{dt}{\sqrt{2t}}$$

$$= \frac{2\sigma}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-t} dt$$

Scanned with CamScanner

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

$$M.D = \frac{2\sigma}{\sqrt{2\pi}} \left[e^{-t}\right]_0^{\infty}$$

$$= \frac{2\sigma}{\sqrt{2\pi}} \left[e^{-\infty} + e^{-0}\right] = \frac{2\sigma}{\sqrt{2\pi}} [0+1]$$

$$M.D = \sigma\sqrt{\frac{2}{\pi}} \text{ (or) } M.D = \frac{4}{5}\sigma$$

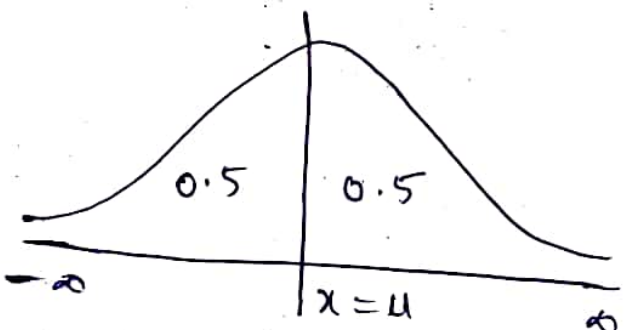
6|9|19

characteristics of N.D

→ The graph of the N.D $f(x)=y$ in the XY- plane is known as the normal curve

$$y = f(x)$$

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



→ The curve is bell shaped curve and symmetrical w.r.t mean i.e about the line $x = \mu$

→ The total area under the normal curve is unity i.e 0.5, 0.5
(↔ x-axis does not t)

→ The curve does not touches the x-axis

→ Mean, Median and Mode of the distribution coincide at $x = \mu$ as the

Scanned with CamScanner

distribution is symmetrical.

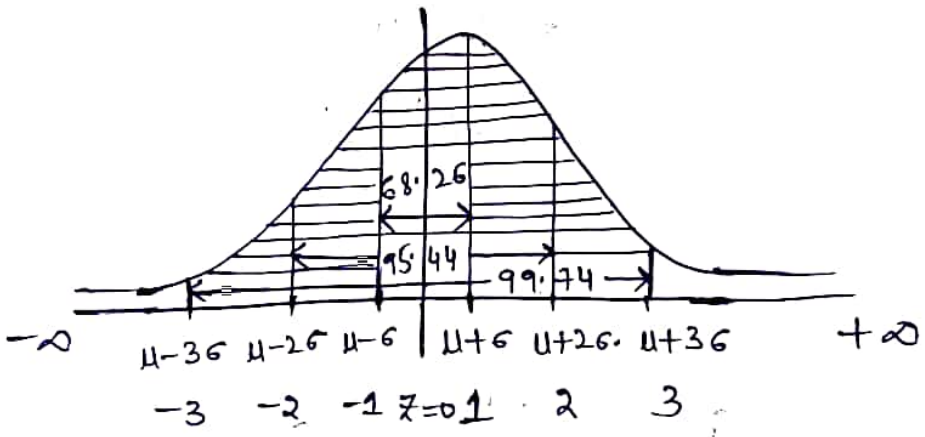
→ The probability that normal variate x with mean (μ) and S.D (σ) lies b/w x_1 and x_2 is given by,

$$P(x_1 \le x \le x_2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

→ The area under the curve is distributed as follows

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

where, z is standard normal variate



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

if, $x = \mu \Rightarrow z = 0$

$$x = \mu - \sigma \Rightarrow z = \frac{\mu-\sigma-\mu}{\sigma} = -1$$

$$x = \mu + \sigma \Rightarrow z = \frac{\mu+\sigma-\mu}{\sigma} = 1$$

$$x = \mu - 2\sigma \Rightarrow z = \frac{\mu-2\sigma-\mu}{\sigma} = -2$$

$$x = \mu + 2\sigma \Rightarrow z = \frac{\mu+2\sigma-\mu}{\sigma} = 2$$

Scanned with CamScanner

$x = \mu - 3\sigma \Rightarrow \frac{\mu - 3\sigma - \mu}{\sigma} = -3$

$x = \mu + 3\sigma \Rightarrow \frac{\mu + 3\sigma - \mu}{\sigma} = 3$

$P(\mu - \sigma \leq x \leq \mu + \sigma) \Rightarrow P(-1 \leq Z \leq 1)$

$= 2P(0 \leq Z \leq 1)$

$= 2A|Z_1|$

$= 2A|1|$

$= 68.26$

$P(\mu - 2\sigma \leq x \leq \mu + 2\sigma) \Rightarrow P(-2 \leq Z \leq 2)$

$= 2P(0 \leq Z \leq 2)$

$= 2A|Z_2|$

$= 2A|2|$

$= 95.44$

$P(\mu - 3\sigma \leq x \leq \mu + 3\sigma) \Rightarrow P(-3 \leq Z \leq 3)$

$= 2P(0 \leq Z \leq 3)$

$= 2A|Z_3|$

$= 2A|3|$

$= 99.74$

Standard Normal variate (Z)

A variable X follows the N.D with mean(μ) & S.D(σ) then the standard Normal variate (Z) is defined as

Scanned with CamScanner

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

Q. To Find the probability density of Normal curve

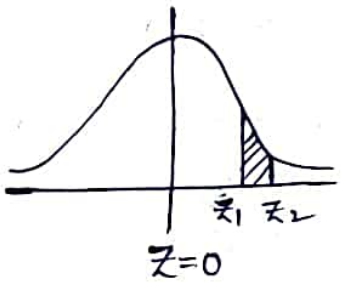
→ The probability that the normal variate x with mean μ and S.D σ lies b/w two specific values x_1 and x_2 with $x_1 < x_2$ can be obtained by using normal curve i.e $P(x_1 < x \leq x_2)$ with $x_1 < x_2$

→ To find z_1, z_2 corresponding to x_1, x_2

i.e $P(x_1 \leq x \leq x_2) = P(z_1 \leq z \leq z_2)$

case (i)

If $z_1, z_2 \geq 0$ and condition $z_1 < z_2$



$P(z_1 \leq z \leq z_2)$

$= P(0 \leq z \leq z_2) - P(0 \leq z < z_1)$

$= A|z_2| - A|z_1|$

= (Area of 0 to z_2) − (Area of 0 to z_1)

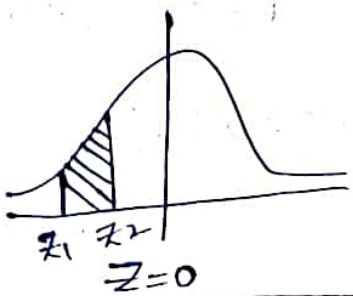
case(ii)

If $z_1, z_2 \leq 0$, and $z_1 < z_2$

$P(z_1 \leq z \leq z_2) =$

$P(0 z_1 \leq z \leq 0) + -$

$P(z_2 \leq z \leq 0)$



Scanned with CamScanner

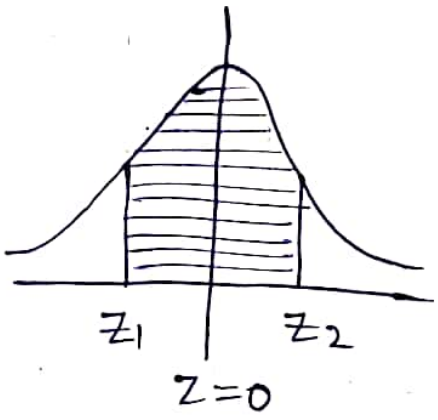
$= A|z_1| - A|z_2|$

case (iii)

If $z_1 < 0$ and $z_2 > 0$

$P(z_1 \le z \le z_2)$

$= P(z_1 \le z \le 0) +$
$P(0 \le z \le z_2)$



$= A|z_1| + A|z_2|$

(Area of z_1 to 0) + (Area of 0 to z_2)

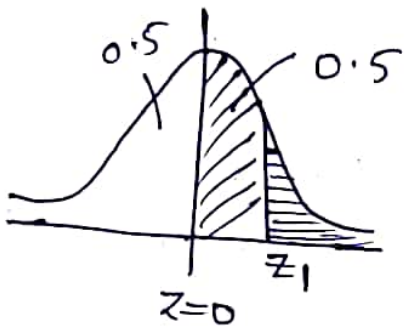
case (iv)

If $P(x \ge x_1) = P(z \ge z_1)$

If $z_1 > 0$

$= 0.5 - P(0 \le z \le z_1)$

$z_1 =$ 



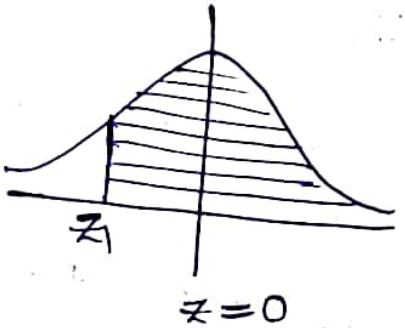
$P(z \ge z_1) = 0.5 - A|z_1|$

If $z_1 < 0$

$P(z \ge z_1)$

$= P(z_1 \le z \le 0) + 0.5$

$= 0.5 + A|z_1|$



Scanned with CamScanner

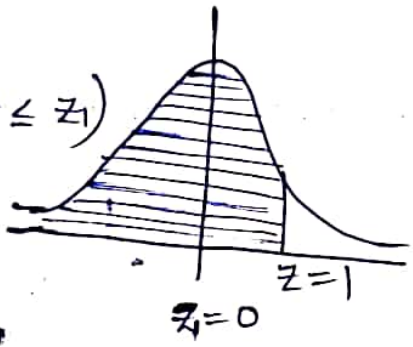
case(v)

$P(X \le x_1) = P(Z \le z_1)$

if $z_1 > 0$

$P(z_0 \le z_1) = 0.5 + P(0 \le z \le z_1)$

$= 0.5 + A|z_1|$

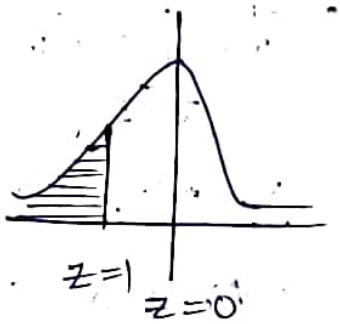


if $z_1 < 0$

$P(z \le z_1) =$

$= 0.5 - P(z_1 \le z \le 0)$

$= 0.5 - A|z_1|$



13/9/19

Q: The mean and standard deviation of the marks obtained by 1000 students in an examination hall 34.5 and 16.5 resp. Find the approximate no. of students expected to obtain marks b/w 30 and 60.

Sol: Given that,

Mean, $\mu = 34.5$

S.D, $\sigma = 16.5$

$n = 1000$

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

To find $P(z_1 \le z \le z_2)$

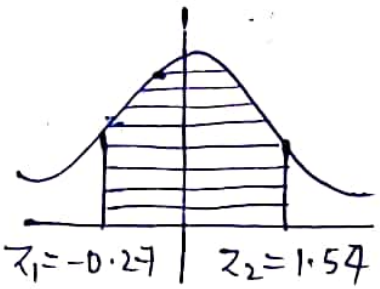
i.e $P(30 \le z \le 60)$

Scanned with CamScanner

$x_1 = 30 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma} = \frac{30 - 34.5}{16.5} = -0.27$

$x_2 = 60 \Rightarrow z_2 = \frac{x_2 - \mu}{\sigma} = \frac{60 - 34.5}{16.5} = 1.54$

$z_1 = -0.27 < 0 \; ; \quad z_2 = 1.54 > 0$



$P(z_1 \leq z \leq z_2) = P(z_1 \leq z \leq 0) + P(0 \leq z \leq z_2)$

$= \text{Area of } z_1 + \text{Area of } z_2$

$= A|z_1| + A|z_2|$

$= A|-0.27| + A|1.54|$

$= A|0.27| + A|1.54|$

$= 0.1064 + 0.4382$

$= 0.5446$

The no. of students who get marks between 30 and 60 $= 0.5446 \times 1000$

$= 5446.$

Q: 1000 students has written an exam the mean is 35 and S.D is 5. Assuming the distribution is normal

(i) How many students to marks lies between 25 and 40.

(ii) How many get more than 40:

Scanned with CamScanner

iii) How many students gets below 20.

Sol: Given that

Mean, $\mu = 35$

S.D, $\sigma = 5$

$n = 1000$

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

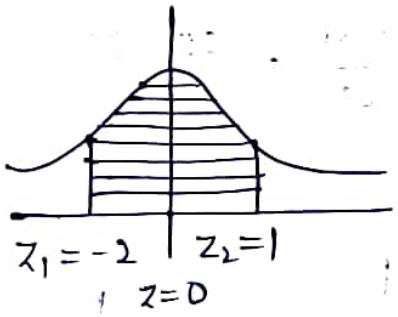
(i) To Find $P(z_1 \leq z \leq z_2)$.

i.e., $P(25 \leq z \leq 40)$

$$x_1 = 25 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma} = \frac{25 - 35}{5} = -2$$

$$x_2 = 40 \Rightarrow z_2 = \frac{40 - 35}{5} = 1$$

$z_1 = -2 < 0$; $z_2 = 1 > 0$



$$P(z_1 \leq z \leq z_2) = P(z_1 \leq z \leq 0) + P(0 \leq z \leq z_2)$$

$$= \text{Area of } z_1 + \text{Area of } z_2$$

$$= A|z_1| + A|z_2|$$

$$= A|-2| + A|1|$$

$$= A|2| + A|1|$$

$$= 0.4772 + 0.3413$$

$$= 0.8185$$

The no. of students who got marks

Scanned with CamScanner

b/w 25 and 40 $= 0.8185 \times 1000$

$= 818.5$

(ii) $P(z > 40) : P(X > 40)$

$x_1 = 40 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma} = \frac{40-35}{5} = 1$

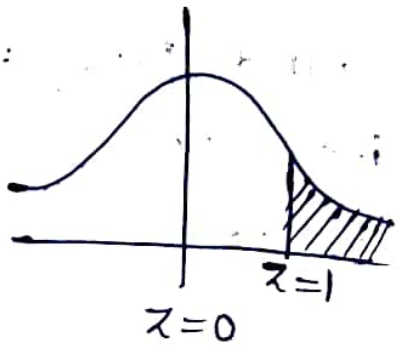
$P(x \geq x_1) = P(z \geq z_1)$

$= 0.5 - A|z_1|$

$= 0.5 - A|1|$

$= 0.5 - 0.3413$

$= 0.1587$



$P(X > 40) = n \times 0.1587 = 1000 \times 0.1587$

$= 158.7$

(iii) $P(z < 20) = P(X < 20)$

$x_1 = 20 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma} = \frac{20-35}{5} = -3$

$P(X \leq x_1) = P(z \leq z_1)$

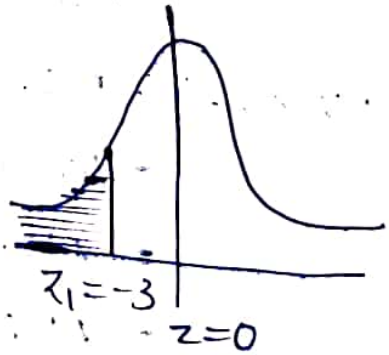
$= 0.5 - A|z_1|$

$= 0.5 - A|-3|$

$= 0.5 - A|3|$

$= 0.5 - 0.4987$

$= 0.0013$



$P(X < 20) = n \times 0.0013$

$= 1000 \times 0.0013$

$= 1.3$

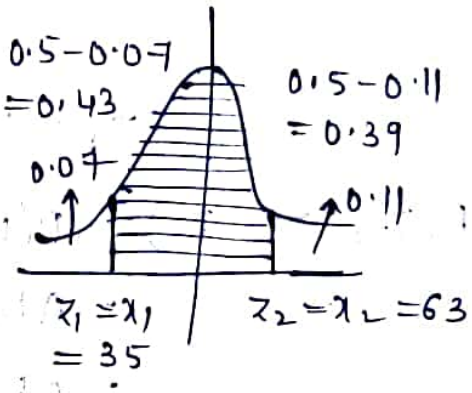
Scanned with CamScanner

Q: Find the mean and standard deviation of the normal distribution in which 7% of the items are under 35 and 89% are under 63.

Sol: Given that;

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

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



$x_1 = 35 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma}$

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

$35 - \mu = -1.48\sigma$

$\mu - 1.48\sigma = 35$ — ①

$x_2 = 63 \Rightarrow z_2 = \frac{x_2 - \mu}{\sigma}$

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

$\sigma \times 1.23 = 63 - \mu$

$1.23\sigma + \mu = 63$ — ②

From ① & ②

$\mu = 50.29, \sigma = 10.33$

Q: If the mean = 70, S.D. = 16. Find

i) $P(38 \leq x \leq 46)$

ii) $P(62 \leq x \leq 86)$

iii) $P(82 \leq x \leq 94)$

Sol (i) $P(38 \leq x \leq 46)$

$x_1 = 38 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma}$

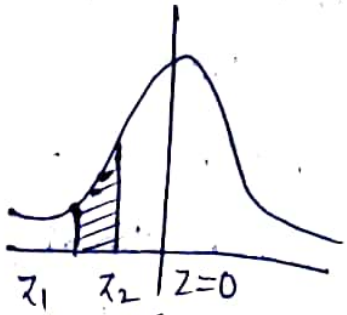
Scanned with CamScanner

$$= \frac{38-70}{16}$$

$$z_1 = -2 < 0$$

$$x_2 = 46 \Rightarrow z_2 = \frac{x_2 - \mu}{\sigma}$$

$$= \frac{46-70}{16}$$



$$z_2 = -1.5 < 0$$

$$P(z_1 \leq z \leq z_2) = P(0 \leq z \leq z_1) - P(0 \leq z \leq z_2)$$

$$= A|z_1| - A|z_2|$$

$$= A|-2| - A|-1.5|$$

$$= A|-2| - A|1.5|$$

$$= 0.4772 - 0.4332$$

$$= 0.044$$

(ii) $P(62 \leq x \leq 86)$

$$x_1 = 62 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma} = \frac{62-70}{16} = -0.5 < 0$$

$$x_2 = 86 \Rightarrow z_2 = \frac{x_2 - \mu}{\sigma} = \frac{86-70}{16} = 1 > 0$$

$$z_1 = -0.5 < 0 \; ; \; z_2 = 1 > 0$$

$$P(z_1 \leq z \leq z_2) = P(z_1 \leq z \leq 0) + P(0 \leq z \leq z_2)$$

$$= A|z_1| + A|z_2|$$

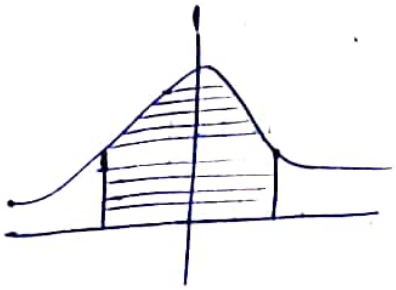
$$= A|-0.5| + A|1|$$

$$= A|0.5| + A|1|$$

$$= 0.1916 + 0.3413$$

Scanned with CamScanner

$= 0.5329$



(iii) $P(82 \leq x \leq 94)$

$$x_1 = 82 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma} = \frac{82 - 70}{16}$$

$$z_1 = 0.75 > 0$$

$$x_2 = 94$$

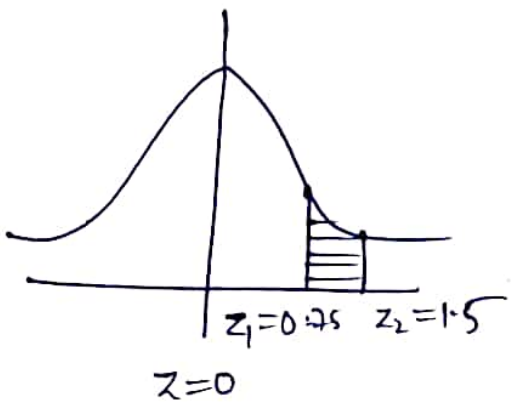
$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{94 - 70}{16}$$

$$= 1.5$$

$$z_2 > 0$$

$z_1 = 0.75 > 0$; $z_2 = 1.5 > 0$

$$z_1 < z_2$$



Scanned with CamScanner

$$P(z_1 \le z \le z_2) = P(0 \le z \le z_2) - P(0 \le z \le z_1)$$

$$= A|z_2| - A|z_1|$$

$$= A|1.5| - A|0.75|$$

$$= 0.4332 - 0.2734$$

$$= 0.1598$$

20/9/19

Scanned with CamScanner

UNIT-III 20/9/19

BASIC STATISTICS.

Curve Fitting

Through curve fitting we learn how to fit linear and non linear equations to the given data by using least square approxim-ations which minimise the sum of the squares of errors (or) deviations.

Fitting of the straight line by the method of least square.

The method of the least square says that the sum of the squares of the errors are minimum and by using this method we find the best fit to the given data.

Let the equation of straight line $y = a+bx$

$$y = a+bx \quad \text{—} ①$$

The normal equations of the straight line

$\Sigma y = na + b\Sigma x$ and $\Sigma xy = a\Sigma x + b\Sigma x^2$

—② —③

(or)

$$y_i = a + bx_i \quad \text{—} ①$$

Then normal equations are

$\Sigma b_i = na + b\Sigma x_i$ and $\Sigma x_i b_i = a\Sigma x_i + b\Sigma x_i^2$

Solve equation ② & ③ we get a & b and substituting the values of a and b in equation we get required straight line

Scanned with CamScanner

Q: Fit a straight line for the following data by using least square method

x	1	2	3	4	6	8
y	2.4	3.1	3.5	4.3	5	6

Sol: The given data

x	1	2	3	4	6	8
y	2.4	3.1	3.5	4.3	5	6

Equation of straight line required

$$y = a+bx \quad \text{—} \quad (1)$$

The two normal equations of the given straight lines are

$$\Sigma y = na + b\Sigma x \quad \text{—} \quad (2)$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \quad \text{—} \quad (3)$$

$$n = 6$$

x	y	x^2	xy
1	2.4	1	2.4
2	3.1	4	6.2
3	3.5	9	10.5
4	4.3	16	17.2
6	5	36	30
8	6	64	48

$\Sigma x = 24 \quad \Sigma y = 24.3 \quad \Sigma x^2 = 130 \quad \Sigma xy = 114.3$

Scanned with CamScanner

substituting these values in eqn ② & ③

eqn ② ⇒ $24.3 = 6 \times a + b \times 24$

$6a + 24b = 24.3$ — ④

eqn ③ ⇒ $114.3 = a \times 24 + b\,130$

$24a + 130b = 114.3$ — ⑤

equating ④ & ⑤

eqn ④ × 4

$24a + 96b = 97.2$

$24a + 130b = 114.3$

$-34b = -17.1$

$b = 0.50294$

we get $a = 2.03823$

sub a & b in eqn ①

$y = 2.03823 + 0.50294x$

Q: Fit a straight line to the following data.

x_i	1	2	3	4	5	6	7
y_i	2.5	5.2	9.7	16.5	29.4	35.5	54.4

Sol: The given data

x_i	1	2	3	4	5	6	7
y_i	2.5	5.2	9.7	16.5	29.4	35.5	54.4

Scanned with CamScanner

The required straight line eqn.

$$y_i = a + bx_i \quad \text{— ①}$$

The 2 normal equations of given st. line

$$\varepsilon y_i = na + b\varepsilon x_i \quad \text{— ②}$$

$$\varepsilon x_i b_i = a\varepsilon x_i + b\varepsilon x_i^2 \quad \text{— ③}$$

$n = 7$

x_i	y_i	x_i^2	$x_i y_i$
1	2.5	1	2.5
2	5.2	4	10.4
3	9.7	9	29.1
4	16.5	16	66.0
5	29.4	25	147
6	35.5	36	213.
7	54.4	49	380.8

$\varepsilon x_i = 28$ $\varepsilon y_i = 153.2$, $\varepsilon x_i^2 = 140$. $\varepsilon x_i y_i = 848.8$

eqn ② $\Rightarrow$ $153.2 = 7 \times a + b \times 28$

$$7a + 28b = 153.2 \quad \text{— ④}$$

eqn ③ $\Rightarrow$ $848.8 = a \times 28 + b \times 140$

$$28a + 140b = 848.8 \quad \text{— ⑤}$$

Solving eqn ④ and ⑤. we get

$$a = -11.82857$$

$$b = 8.42857$$

eqn ① $\Rightarrow$ $y_i = -11.82857 + 8.42857 x_i$

Scanned with CamScanner

Q: Fit a straight line to the following data

x_i	20	30	40	50	60	70
y_i	800·3	800·4	800·6	800·7	800·9	801

sol:

The required straight line eqn

$$y_i = a + bx_i \quad \text{—— (1)}$$

The 2 normal equations of given st. line

$$\Sigma y_i = na + b\Sigma x_i \quad \text{— (2)}$$

$$\Sigma x_i y_i = a\Sigma x_i + b\Sigma x_i^2 \quad \text{— (3)}$$

$$n = 6$$

x_i	y_i	x_i^2	$x_i y_i$
20	800·3	400	16006
30	800·4	900	24012
40	800·6	1600	32024
50	800·7	2500	40035
60	800·9	3600	48054
70	801	4900	56070

$\Sigma x_i = 270$

$\Sigma y_i = 4803.9$

$\Sigma x_i^2 = 13900$

$\Sigma x_i y_i = 216201$

Scanned with CamScanner

eqn ② $\Rightarrow 4803.9 = 6a + b \times 270$

$6a + 270b = 4803.9$ — ④

eqn ③ $\Rightarrow 216201 = a \times 270 + b \times 13900$

$270a + 13900b = 216201$ — ⑤

solving eqn ④ & ⑤ we get

$a = 344.25$,

$b = 8.86696$

$y_i = a + bx_i$

$y_i = 344.25 + 8.86696 x_i$

Fitting of second degree polynomial or parabola by the method of least square.

let the equation of parabola or polynomial by the method of least squares

$$y = a + bx + cx^2 \quad \text{—} ①$$

The normal equations of parabola or polynomial are

$$\Sigma y = na + b\Sigma x + c\Sigma x^2 \quad \text{—} ②$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3 \quad \text{—} ③$$

$$\Sigma x^2 y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4 \quad \text{—} ④$$

Solve ②, ③, ④ we get a, b, c values and these values sub. in eqn ① we get required polynomial or parabola equation

Q: Fit a second degree polynomial by the following data by using least square method.

Scanned with CamScanner

x	2	4	6	8	10
y	3.07	12.85	31.47	57.38	91.29

Parabola

Sol: The required st. line eqn is

$$y = a + bx + cx^2 \quad \text{—— (1)}$$

The 3 normal equations of parabola are

$$\varepsilon y = na + b\varepsilon x + c\varepsilon x^2 \quad \text{—— (2)}$$

$$\varepsilon xy = a\varepsilon x + b\varepsilon x^2 + c\varepsilon x^3 \quad \text{—— (3)}$$

$$\varepsilon x^2 y = a\varepsilon x^2 + b\varepsilon x^3 + c\varepsilon x^4 \quad \text{—— (4)}$$

$$n = 5$$

x	y	x^2	x^3	x^4	xy	x^2y
2	3.07	4	8	16	6.14	12.28
4	12.85	16	64	256	51.4	205.6
6	31.47	36	216	1296	188.84	1132.92
8	57.38	64	512	4096	459.04	3672.32
10	91.29	100	1000	10,000	912.9	9129

$\varepsilon x = 30$

$\varepsilon y = 196.06$

$\varepsilon x^2 = 220$

$\varepsilon x^3 = 1800$

$\varepsilon x^4 = 15664$

$\varepsilon xy = 1618.3$

$\varepsilon x^2 y = 14152.1$

Scanned with CamScanner

Sub these values in eqn 2, 3, 4.

eqn ② ⇒ $19606 = 5a + b \times 30 + c \times 220$

$5a + 30b + 220c = 19606$ — ⑤

eqn ③ ⇒ $1618.3 = a \times 30 + b \times 220 + c \times 1800$

$30a + 220b + 1800c = 1618.3$ — ⑥

eqn ④ ⇒ $14152.1 = a \times 220 + b \times 1800 + c \times 15664$

$220a + 1800b + 15664c = 14152.1$ — ⑦

Solving 5, 6, 7.

we get

$a = 0.6935$

$b = -0.854$,

$c = 0.9918$

Sub a, b, c in eqn ①

$y = a + bx + cx^2$

$y = 0.6935 - 0.854x + 0.9918x^2$

24/9/19

Q: Fit a second degree polynomial for the following data.

x	1	2	3	4	5	6	7
y	2.3	5.2	9.7	16.5	29.4	35.5	54.4

Sol: The required eqn of polynomial is

$y = a + bx + cx^2$ — ①

Scanned with CamScanner

The 3 normal eqn's of polynomial are

$$\Sigma y = na + b\Sigma x + c\Sigma x^2 \quad \text{— (2)}$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3 \quad \text{— (3)}$$

$$\Sigma x^2 y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4 \quad \text{— (4)}$$

$n = 7$

x	y	x^2	x^3	x^4	xy	x^2y
1	2.3	1	1	1	2.3	2.3
2	5.2	4	8	16	10.4	20.8
3	9.7	9	27	81	29.1	87.3
4	16.5	16	64	256	66	264
5	29.4	25	125	625	147	735
6	35.5	36	216	1296	213	1278
7	54.4	49	343	2401	3808	2665.6

$\Sigma x = 28$

$\Sigma y = 153$

$\Sigma x^2 = 140$

$\Sigma x^3 = 784$

$\Sigma x^4 = 4676$

$\Sigma xy = 8975.8$ 848.6

$\Sigma x^2 y = 5053$

Scanned with CamScanner

Sub these values in eqn 2, 3, 4.

eqn ② ⇒ $153 = 7a + b(28) + c(140)$

$7a + 28b + 140c = 153$ — ⑤

eqn ③ ⇒ $848.6 = a(28) + b(140) + c(784)$

$28a + 140b + 784c = 848.6$ — ⑥

eqn ④ ⇒ $5053 = a(140) + b(784) + c(4676)$

$140a + 784b + 4676c = 5053$ — ⑦

Solving 5, 6, 7 we get.

$a = 2.371$

$b = -1.092$

$c = 1.192$

sub a, b, c values in eqn ①

$y = a + bx + cx^2$

$y = 2.371 - 1.092x + 1.192x^2$

27/9/19

Q: Fit a st. line by method of least squares.

(1)

x	1	2	3	4	5
y	14	27	40	55	68

(2)

x	6	7	7	8	8	8	9	9	10
y	5	5	4	5	4	3	4	3	3

Q: Fit a second degree polynomial equation.

Scanned with CamScanner

(1)

x	0	1	2	3	4
y	1	1.8	1.3	2.5	6.3

(2) Fit a $y = a + bx + cx^2$

x	0	1.0	2.0
y	1.0	6.0	17.0

Fitting of an exponential curve of the form $y = ae^{bx}$

Consider the equation of exponential curve $y = ae^{bx}$. To convert this equation into linear form take log on both sides

$$y = ae^{bx}$$

$$\log y = \log(ae^{bx})$$

$$\log y = \log a + \log e^{bx}$$

$$\log y = \log a + bx \log e$$

$$\log y = \log a + bx$$

let, $\log_e y = Y$; $\log_e a = A \Rightarrow a = e^A$

$$Y = A + bx \longrightarrow (1)$$

∴ The normal eqn is of $Y = A + bx$ are

$$\Sigma Y = nA + b\Sigma x \longrightarrow (2)$$

$$\Sigma xY = A\Sigma x + b\Sigma x^2 \longrightarrow (3)$$

Solving eq (2) and (3) we get, A, b.

Q: Fit a curve of the form $y = ae^{bx}$ from the following data

Scanned with CamScanner

x	77	100	185	239	285
y	2.4	3.4	7	11.1	19.6

Sol: The required equation of exponential curve is

$$y = ae^{bx} \longrightarrow (1)$$

Applying log on both sides.

$$\log y = \log(ae^{bx})$$

$$\log y = \log a + \log e^{bx}$$

$$\log y = \log a + bx \log e$$

$$\log y = \log a + bx$$

Let, $\log_e y = Y$; $\log_e a = A \Rightarrow a = e^A$

$$Y = A + bx \longrightarrow (2)$$

The normal eqn's of the curve are.

$$\Sigma Y = nA + b\Sigma x \longrightarrow (3)$$

$$\Sigma xY = A\Sigma x + b\Sigma x^2 \longrightarrow (4)$$

$$n = 5$$

x	y	Y	x^2	xY
77	2.4	0.380	5929	29.26
100	3.4	0.531	10000	53.1
185	7	0.845	34225	156.325
239	11.1	1.045	52441	239.305
285	19.6	1.292	81225	368.22

$\Sigma x = 876$

$\Sigma Y = 4.093$

Scanned with CamScanner

x	y	Y	x^2	xy
77	2.4	0.875	5929	67.375
100	3.4	1.223	10000	122.3
185	7	1.945	34225	359.825
239	11.1	2.406	57121	575.034
285	19.6	2.975	81225	847.875

$\Sigma x = 886$

$\Sigma Y = 9.424$

$\Sigma x^2 = 188500$

$\Sigma xY = 1972.409$

Sub. this values in eqn ③ & ④

③ $\Rightarrow 9.424 = 5A + b(886)$

$5A + 886b = 9.424 \longrightarrow ⑤$

④ $\Rightarrow 1972.409 = A(886) + b(183500)$

$886A + 183500b = 1972.409 \longrightarrow ⑥$

Solving eq ⑤ and ⑥

$A = 0.1832$

$b = 9.602 \times 10^{-3}$

$a = e^A = e^{0.1832} = 1.201$

Sub. a, b values in eqn ①

$y = ae^{bx}$

$y = 1.201e^{(9.602 \times 10^{-3})x}$.

Scanned with CamScanner

Q: Fit a curve $y = ae^{bx}$ from following data

x	2	3	4	5	6
y	144	172.8	207.4	248.5	298.5

Sol: The required exponential curve

$$y = ae^{bx} \quad \text{---} ①$$

Applying log on B.S

$$\log y = \log(ae^{bx})$$

$$\log y = \log a + \log e^{bx}$$

$$\log y = \log a + bx \log e$$

$$\log y = \log a + bx$$

let, $\log_e y = Y$; $\log_e a = A \Rightarrow a = e^A$

$$Y = A + bx \quad \text{---} ②$$

The normal equations of the curve are

$$\Sigma Y = nA + b\Sigma x \quad \text{---} ③$$

$$\Sigma xY = A\Sigma x + b\Sigma x^2 \quad \text{---} ④$$

x	y	x^2	Y	xY
77	2.4	5929	0.875	67.375
100	3.4	10000	1.223	122.3
185	7	34225	1.945	359.825
239	11.1	57121	2.406	575.034
285	19.6	81225	2.975	847.871

$\Sigma x = 886$

$\Sigma Y = 9.424$

$\Sigma xY = 1972.409$

Scanned with CamScanner

x	y	Y	x^2	xY
2	144	4.969	4	9.938
3	172.8	5.152	9	15.456
4	207.4	5.334	16	21.336
5	248.5	5.515	25	27.575
6	298.5	5.698	36	34.188

$\Sigma x = 20$

$\Sigma Y = 26.668$

$\Sigma x^2 = 90$

$\Sigma xY = 108.493$

Sub this values in ③ & ④

③ $\Rightarrow 26.668 = 5A + b(20)$

$5A + 20b = 26.668$ — ⑤

④ $\Rightarrow 108.493 = A(20) + b(90)$

$20A + 90b = 108.493$ — ⑥

Solving ⑤ & ⑥ we get

$A = 4.605$

$b = 0.182$

$a = e^A = e^{4.605} = 99.982$

Sub a, b in eqn ①

$y = ae^{bx}$

$y = 99.982^{(80.182)x}$

$y = 99.982^{0.182x}$.

Scanned with CamScanner

POWER CURVES.

There are two types of power curves.

1. $y = ab^x$

Apply log on B.S

$$\log y = \log(ab^x)$$
$$\log y = \log a + \log b^x$$
$$\log y = \log a + x \log b$$

Let, $\log y = Y$; $\log a = A \Rightarrow a = e^A$

$\log b = B \Rightarrow b = e^B$

$$Y = A + Bx \quad \text{---} (1)$$

The normal eqn's of $Y = A + Bx$ are

$$\Sigma Y = nA + B\Sigma x \quad \text{---} (2)$$
$$\Sigma xY = A\Sigma x + B\Sigma x^2 \quad \text{---} (3)$$

Solving eqn (2) & (3) we get A, B

2. $y = ax^b$

Apply log on B.S

$$\log y = \log(ax^b)$$
$$\log y = b \log ax^b$$
$$\log y = \log a + \log x^b$$

Let, $\log y = Y$; $\log a = A \Rightarrow a = e^A$

$$\log x = X$$
$$Y = A + bX \quad \text{---} (1)$$

The normal eqn's of $Y = A + bX$ are

$$\Sigma Y = nA + b\Sigma X \quad \text{---} (2)$$
$$\Sigma XY = A\Sigma X + b\Sigma x^2 \quad \text{---} (3)$$

Solving eqn (2) and (3) we get A, b

Scanned with CamScanner

Q: Fit a curve of the form $y = ab^x$ to the following data.

x	1	2	3	4
y	7	11	17	27

Sol: The required eqn. of curve is

$$y = ab^x \quad \text{—— (1)}$$

$$\log y = \log(ab^x)$$

$$\log y = \log a + \log b^x$$

$$\log y = \log a + x \log b$$

let,

$$\log y = Y ; \log a = A \Rightarrow a = e^A$$

$$\log b = B \Rightarrow b = e^B$$

$$Y = A + Bx \quad \text{—— (2)}$$

The normal eqn's of the curve are

$$\Sigma Y = nA + B\Sigma x \quad \text{—— (3)}$$

$$\Sigma xY = A\Sigma x + B\Sigma x^2 \quad \text{—— (4)}$$

x	y	Y	x^2	xY
1	7	1.945	1	1.945
2	11	2.397	4	4.794
3	17	2.833	9	8.499
4	27	3.295	16	13.18

Scanned with CamScanner

$\Sigma x = 10$

$\Sigma Y = 10.47$

$\Sigma x^2 = 30$

$\Sigma xY = 28.418$

Sub this values in ③ and ④

③ $\Rightarrow 10.47 = 4A + B(10)$

$4A + 10B = 10.47$ — ⑤

④ $\Rightarrow 28.418 = A(10) + B(30)$

$10A + 30B = 28.418$ — ⑥

Solving ⑤ and ⑥ we get

$A = 1.496$

$B = 0.448$

$a = e^A = e^{1.496} = 4.463$

$b = e^B = e^{0.448} = 1.565$

Sub a, b values in eqn ①

$y = ab^x$

$y = 4.463(1.565)^x$

Scanned with CamScanner

UNIT - III

22/10/19

CORRELATION AND REGRESION

→ Previously we discussed the characterist-ics of only one variable (i.e marks, height, weight etc). This type of analysis is called univariant analysis

→ Now, we will discuss the relation b/w two variables, this analysis is called byvariant analysis.

→ Correlation discuss the relation b/w two or more variants.

Scanned with CamScanner

CORRELATION.

Correlation is a statistical analysis which measure and analyse withe two variable fluctuate with reference to each other.

Types of Correlation.

1. Positive and negative relation.
2. Partial and total relation.
3. Simple and multiple relation.
4. Linear and non linear relation.

1. POSITIVE AND NEGATIVE CORRELATION:

It depends on how the two variables are fluctuated,

(i) If the two variables tends to move together in same direction (↑↑ or ↓↓) i.e both variables are ↑ or ↓ then there is a positive correlation b|w two variables.

(ii) If the two variables tends to move opposite direction (↑↓ (or) ↓↑). i.e: one is increasing and other is decreasing, then there is a -ve correlation b|w two variables.

2. PARTIAL AND TOTAL CORRELATION.

(i) If we study the relation between

Scanned with CamScanner

2 variables excluding some other variab -les is called a partial correlation.

EX: The relation b|w sales and price excluding demand.

→ If we study the relation b|w all the variables is called total correlation.

EX: Relation b|w sales, price and demand.

3. SIMPLE AND MUILTIPLE CORRELATION.

→ If we study the relation b|w only two variables is called simple correlation.

→ The relation b|w more than 2 variables is called a multiple correlation.

4. LINEAR AND NON-LINEAR CORRELATION.

→ Suppose we ploted the 2 points of 2 variables on a graph which gives some linear equation (or) the relation b|w 2 variables is uniform then there is a linear correlation b|w 2 variables. otherwise there is a non linear correlation.

KARL PERSON'S CORRELATION COEFFICIENT:

→ It is also known as product moment correlation

$$r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}}$$

Scanned with CamScanner

Where, $X = x - \bar{x}$

$Y = y - \bar{y}$

Properties of correlation coefficient

1) The limits for the correlation coefficient are $-1 \leq r \leq 1$ i.e the correlation coefficient cannot exceed the 1 numerical value $|r| \leq 1$

2) If $r = 1$ then the correlation is perfect and positive.

3) If $r = -1$ then the correlation is perfect and negative.

4) If $r = 0$ then there is no relation b/w two variables.

Q: Find the karl person's correlation coefficients b/w heights and weights given below

Hts in inch	57	59	62	63	64	65	55	58	57
wts in grms	113	117	126	126	130	129	111	116	112

Sol:

$$r = \frac{\Sigma XY}{\sqrt{\Sigma x^2 \Sigma y^2}}$$

Scanned with CamScanner

x	y	$X = x - \bar{x}$	$Y = y - \bar{y}$	XY	X^2	Y^2
57	113	-3	-7	21	9	49
59	117	-1	-3	3	1	6
62	126	2	6	12	4	36
63	126	3	6	18	9	36
64	130	4	10	40	16	100
65	129	5	9	45	25	81
55	111	-5	-9	45	25	81
58	116	-2	-4	8	4	16
57	112	-3	-8	24	9	64

$$\bar{x} = \frac{\Sigma x}{n} = 60 \qquad \Sigma XY = 216$$

$$\bar{y} = \frac{\Sigma y}{n} = 120 \qquad \Sigma X^2 = 102$$

$$\Sigma Y^2 = 472$$

$$r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}}$$

$$= \frac{216}{\sqrt{102 \times 472}}$$

$$r = 0.98$$

Q: Findout the coefficient of correlation in the following case

Ht. of father	65	66	67	67	68	69	71	73
Wt. of son	67	68	64	68	72	70	69	70

Scanned with CamScanner

Sol:

x	y	$X = x - \bar{x}$	$Y = y - \bar{y}$	XY	X^2	Y^2
65	67	−3	−2	6	9	4
66	68	−2	−1	2	4	1
67	64	−1	−5	5	1	25
67	68	−1	−1	1	1	1
68	72	0	3	0	0	9
69	70	1	1	1	1	1
71	69	3	0	0	9	0
73	70	5	1	5	25	1

$$\bar{x} = \frac{\Sigma x}{n} = 68.25 \simeq 68 \qquad \Sigma X = 2$$

$$\bar{y} = \frac{\Sigma y}{n} = 68.5 \simeq 69 \qquad \Sigma Y = -4$$

$$\Sigma XY = 20$$

$$\Sigma X^2 = 50$$

$$\Sigma Y^2 = 42$$

The correlation coefficient,

$$r = \frac{\Sigma XY - \frac{\Sigma X \Sigma Y}{n}}{\sqrt{\Sigma X^2 - \frac{(\Sigma X)^2}{n}}\sqrt{\Sigma Y^2 - \frac{(\Sigma Y)^2}{n}}}$$

$$= \frac{20 - \frac{(2)(-4)}{8}}{\sqrt{50 - \frac{(2)^2}{8}}\sqrt{42 - \frac{(-4)^2}{8}}}$$

Scanned with CamScanner

$$= \frac{21}{\sqrt{\frac{99}{2}}\sqrt{40}}$$

$$r = 0.471$$

RANK CORRELATION COEFFICIENT

Charle's Spearman found the method of correlation coefficient by rank and is defined as

$$\text{Rank, } P = 1 - \frac{6\Sigma D^2}{N(N^2-1)}$$

where,

D = difference b/w two ranks.

N = no. of observations

NOTE:

Equal or repeated ranks.

→ If in the given data, the ranks are repeated then the ranks correlation coefficient

$$P = 1 - \frac{6\left[\Sigma D^2 + \frac{1}{12}m(m^2-1) + \frac{1}{12}m(m^2-1) + \cdots\right]}{N(N^2-1)}$$

Where,

m = no. of itemes the ranks were repeated

Scanned with CamScanner

Q: A random sample of 5 college students are selected and their grades in mathematics & statistics are found to be

	1	2	3	4	5
maths	85	60	73	40	90
statistics	93	75	65	50	80

Calculate spearmen's rank correlation coefficient

Sol:

Marks in Maths (x)	Ranks (R_x)	Marks in Stats (y)	Ranks (R_y)	D ($R_x - R_y$)	D^2
85	2	93	1	1	1
60	4	75	3	1	1
73	3	65	4	-1	1
40	5	50	5	0	0
90	1	80	2	-1	1

$$\text{Rank, } \rho = 1 - \frac{6\Sigma D^2}{N(N^2-1)}$$

$$= 1 - \frac{6(4)}{5(25-1)}$$

$$= 1 - \frac{24}{5 \times 24}$$

$$\rho = -0.191$$

Scanned with CamScanner

$$= 1 - \frac{1}{5}$$

$$P = 0.8$$

25/10/19

Q: Obtain the rank correlation coefficient for the following data

x	68	64	75	50	64	80	75	40	55	64
y	62	58	68	45	81	60	68	48	50	70

Sol:

	x	y	R_x	R_y	$D = R_x - R_y$	D^2
4	68	62	4	5	−1	1
6	64	58	6	7	−1	1
2.5	75	68	2.5	3.5	−1	1
9	50	45	9	10	−1	1
6	64	81	6	1	5	25
1	80	60	1	6	−5	25
2.5	75	68	2.5	3.5	−1	1
10	40	48	10	9	1	1
8	55	50	8	8	0	0
6	64	70	6	2	4	16

$$P = 1 - \frac{6\left[\Sigma D^2 + \frac{1}{12}m(m^2-1) + \frac{1}{12}m(m^2-1) + \dots\right]}{N(N^2-1)}$$

$$\Sigma D^2 = 72$$

$$N = 10$$

Scanned with CamScanner

D = Diff. b/w two ranks. $(R_x - R_y)$

N = No. of observations.

m = No. of items repeated.

$$P = 1 - \frac{6\left[72 + \frac{1}{12}2(2^2-1) + \frac{1}{12}3(3^2-1) + \frac{1}{12}2(2^2-1)\right]}{10(10^2-1)}$$

$$= 1 - 0.454$$

$$P = 0.546.$$

REGRESSION.

The statistical method which helps us to estimate the unknown value of one variable from the known value of the related variable is known as Regression.

Uses of Regression.

→ we can calculate correlation coefficient with the help of the regression coefficient.

→ It is highly valuable in economics and bussincss

→ Widely used for production purpose

Comparision between Regression and correlation.

Scanned with CamScanner

CORRELATION	REGRESSION.
1. The correlation coeff -icient is a measures of degree of co-variability b/w two variables	1. The regression establish a functional relation b/w dependent and independent variables.
2. In this both variables x & y are random variables	2. Here one variable is random, the other one is fixed variable.
3. correlation coefficients is a relative measure.	3. The regression coefficient is a obsolute values

Regression coefficient

Regression equation is a algebraic expansion of the regression line. It can be classified into two types.

1. Regression equation of y on x.
2. Regression equation of x on y

1. REGRESSION EQUATION OF Y ON X

It is given by $y - \bar{y} = b_{yx}(x - \bar{x})$

where, $b_{yx} = r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma xy}{\Sigma x^2}$ (exact mean)

$= r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma XY - \frac{\Sigma X \Sigma Y}{n}}{\Sigma X^2 - \frac{(\Sigma X)^2}{n}}$ (Assumed mean)

Scanned with CamScanner

2. REGRESSION EQUATION OF X ON Y.

(A statistical method)

It is given by $(x-\bar{x}) = b_{xy}(y-\bar{y})$

where,

$$b_{xy} = r\frac{\sigma_x}{\sigma_y} = \frac{\Sigma XY}{\Sigma Y^2} \quad \text{(Exact mean)}$$

$$= r\frac{\sigma_x}{\sigma_y} = \frac{\Sigma XY - \frac{\Sigma X \Sigma Y}{n}}{\Sigma Y^2 - \frac{(\Sigma Y)^2}{n}} \quad \text{(Assumed mean)}$$

NOTE:

The relation b/w correlation coefficient and regression coefficient.

$$r = \sqrt{b_{yx} b_{xy}}$$

Q: Price indices of cotton and wool are given below for the 12 months of the year obtain the equations of lines of regression b/w the indices.

cotton(x)	78	77	85	88	87	82	81	77	76	83	97	93
wool(y)	84	82	82	85	89	90	88	92	83	89	98	99

Scanned with CamScanner

Sol:

x	y	$X = x - \bar{x}$	$Y = y - \bar{y}$	XY	X^2	Y^2
78	84	−6	−4	24	36	16
77	82	−7	−6	42	42	36
85	82	−1	−6	−6	1	36
88	85	4	−3	−12	16	9
87	90	3	1	3	9	1
82	90	−2	2	−4	4	4
81	88	−3	0	0	9	0
77	92	−7	4	−28	49	16
76	83	−8	−5	40	64	25
83	89	−1	−1	1	1	1
97	98	13	10	130	169	100
93	99	9	11	99	81	121

$$\bar{x} = \frac{\Sigma x}{n} = \frac{1004}{12} = 83.66 \simeq 84$$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{1061}{12} = 88.41 \simeq 88$$

$\Sigma XY = 287$, $\Sigma X = -4$

$\Sigma X^2 = 488$, $\Sigma Y = 5$

$\Sigma Y^2 = 365$

Scanned with CamScanner

Regression equation of Y on X

$$b_{yx} = \frac{\Sigma XY - \frac{\Sigma X \Sigma Y}{n}}{\Sigma X^2 - \frac{(\Sigma X)^2}{n}}$$

$$= \frac{287 - (-4)(5)/12}{488 - \frac{(-4)^2}{12}}$$

$$= 0.593.$$

Regression equation of X on Y.

$$b_{xy} = \frac{\Sigma XY - \frac{\Sigma X \Sigma Y}{n}}{\Sigma Y^2 - \frac{(\Sigma Y)^2}{n}}$$

$$= \frac{287 - \frac{(-4)(5)}{12}}{365 - \frac{25}{12}}$$

$$= 0.786$$

The regression relation of Y on x is

$$(y - \bar{y}) = b_{yx}(x - \bar{x})$$

$$(y - 88.41) = 0.593(x - 83.66)$$

$$y - 88.41 = 0.593x - 49.61$$

$$0.593x - y + 38.79962 = 0$$

$$y = 38.79962 + 0.593x.$$

Scanned with CamScanner

The regression equation of x on y is

$$(x-\bar{x}) = b_{xy}(y-\bar{y})$$

$$(x-83.66) = 0.7895(y-88.41)$$

$$x-83.66 = 0.7895y - 70.285$$

$$x = 0.7895y + 13.375.$$

Angle b/w two regression lines.

W.K.T the regression equation of y on x

$$y-\bar{y} = r\frac{\sigma_y}{\sigma_x}(x-\bar{x}) \quad \text{---} \ (1)$$

slope of eqn (1) is $m_1 = r\frac{\sigma_y}{\sigma_x}$

The regression eqn of x on y

$$\text{slope } \left(x-\bar{x} = r\frac{\sigma_x}{\sigma_y}(y-\bar{y})\right) \quad \text{---} \ (2)$$

$$(y-\bar{y}) = \frac{1}{r}\frac{\sigma_y}{\sigma_x}(x-\bar{x}) \quad \text{---} \ (2)$$

slope of eqn (2) is $m_2 = \frac{\sigma_x}{\sigma_y}\frac{1}{r}\frac{\sigma_y}{\sigma_x}$

If m_1 & m_2 are slopes of 2 lines then

$$\tan\theta = \left|\frac{m_1 - m_2}{1+m_1m_2}\right|$$

$$= \frac{r\frac{\sigma_y}{\sigma_x} - \frac{1}{r}\frac{\sigma_y}{\sigma_x}}{1+\left(r\frac{\sigma_y}{\sigma_x}\right)\left(\frac{1}{r}\frac{\sigma_y}{\sigma_x}\right)}$$

$$= \frac{\frac{\sigma_y}{\sigma_x}\left(r-\frac{1}{r}\right)}{1+\left(\frac{\sigma_y}{\sigma_x}\right)^2}$$

Scanned with CamScanner

$$= \frac{\left(\gamma - \frac{1}{\gamma}\right)\frac{\sigma_y}{\sigma_x}}{1+\left(\frac{\sigma_y}{\sigma_x}\right)^2}$$

$$\tan\theta = \frac{\frac{\sigma_y}{\sigma_x}\left(\gamma - \frac{1}{\gamma}\right)}{1+\left(\frac{\sigma_y}{\sigma_x}\right)^2}$$

$$\tan\theta = \frac{\sigma_x \sigma_y}{\sigma_x^2 \sigma_y^2}\left[\gamma - \frac{1}{\gamma}\right]$$

Scanned with CamScanner

UNIT – IV

1/11/19

TESTING OF HYPOTHESIS – I

Population	Notation	Sample	Notation.
Mean	μ	mean	$\bar{x}$
variance	σ^2	Variance	s^2
S.D	σ	S.D	s
Proportion	P	proportion	p
site	N	site	n

NULL HYPOTHESIS.

Q: A sample of 400 items is taken from the population whose S.D is 10 and mean of sample is 40. Test whether the sample has come from a population with mean 38 and also find 95% confidence interval.

Sol: Given,

$n = 400$

$\sigma = 10$

$\bar{x} = 40$ — sample mean

$\mu = 38$

$Z_{\alpha/2} = 9.6$ — Tabulated value.

Step – 1.

Null Hypothesis H_0

The sample has drawn from the

Scanned with CamScanner

population with mean (μ) = 38

$\mu = 38$

$H_0 = \mu = \mu_0$

$\mu = 38$

step 2 : Alternative Hypothesis (H_1)

$H_1 : \mu \neq \mu_0$

$\mu \neq 38$

step 3 : Level of significance

$\alpha = 95\%$

$z_{\alpha/2} = 1.96$

step 4 : Test statistics.

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{40 - 38}{10\sqrt{400}} = 4.$$

Conclusion.

$z_{tabulated\ value} = 1.96$

$z_{calculated\ value} = 4$

$z_{cal} > z_{tab}$

Confidence Interval

$$= \left(\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$$

$$= \left(40 - 1.96 \frac{10}{\sqrt{400}}, 40 + 1.96 \frac{10}{\sqrt{400}}\right)$$

$$= (39.02, 40.98)$$

Scanned with CamScanner

The population mean μ is 38 does not lies in blw the above interval. Hence H_0 is rejected.

2) A sample of 64 students have a mean weight of 70kgs can this sample regarded from the population with mean weight 56 kgs and SD 25kgs.

Sol : Given,

$n = 64$

$\bar{x} = 70$

$\mu = 56$

$\sigma = 25$

step – I

H_0

population with mean(μ) = 56

$\mu = 56$

$H_0 : \mu = \mu_0$

step 2: Alternative Hypothesis (H_1)

$H_1 : \mu \neq \mu_0$

$\mu \neq 56$.

step 3: Level of significance

$\alpha = 95\%$

$Z_{\alpha/2} = 1.96$.

step 4: Test statistics.

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{70 - 56}{25/\sqrt{64}} = 4.48$$

Scanned with CamScanner

conclusion.

z_{tab} val $= 1.96$

z_{cal} val $= 4.48$

$z_{cal} > z_{tab}$.

$\therefore H_0$ is rejected.

TESTING OF HYPOTHESIS: DIFFERENCES BETWEEN TWO MEANS.

Let n_1, n_2 are two sample sizes having means $\bar{x}$ & $\bar{y}$ drawn from two populations with mean μ_1 & μ_2 and S.D σ_1 & σ_2 resp.

PROCEDURE

step 1: Null Hypothesis (H_0)

$H_0 : \mu_1 = \mu_2$

step 2: Alternative hypothesis (H_1)

$H_1 : \mu_1 \neq \mu_2$ ($\mu_1 < \mu_2$, $\mu_1 > \mu_2$)

step 3: Level of significance α

step 4: Test statistics.

$$z = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

conclusion:

$z_{cal} < z_{tab} \Rightarrow H_0$ accepted

$z_{cal} > z_{tab} \Rightarrow H_0$ rejected.

Scanned with CamScanner

NOTE:

If two samples have been drawn from two population with common S.D (σ) then,

$$z = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Q: The mean yield of wheat from a district A was 210 pounds with a S.D 10 pounds per acre from a sample of 100 plots. In another district the mean yield was 220 pounds with a S.D 12 pounds from a sample of 150 plots. Assumed that S.D of yield in entire state was 11 pounds. Test whether if there is any significance blw the mean of yield of crop.

Sol: Given.

$n_1 = 100, \; n_2 = 150$

$S_1 = 10, \; S_2 = 12$

$\sigma = 11$

$\bar{x} = 210, \; \bar{y} = 220$

1) Null Hypothesis $H_0 : \mu_1 = \mu_2$

$H_0 : \mu_1 = \mu_2$

2) Alternate hypothesis

$H_1 : \mu_1 \neq \mu_2$

Scanned with CamScanner

3) level of significance

$$\alpha = 95\%$$

$$z_{\alpha/2} = 1.96$$

4) Test statistics

$$z = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{(210 - 220)}{11 \sqrt{\frac{1}{100} + \frac{1}{150}}} \quad (\because \mu_1 = \mu_2)$$

$$z = -7.041$$

$$|z| = 7.041$$

$z_{cal} > z_{tab}$

H_0 rejected.

* A sample of 900 members has a mean of 3.4cm and S.D of 2.61cm. Is this sample drawn from a large population with mean 3.25 cm and SD 2.61cm If the population is normal and it's mean is unknown then find 95% confidence interval for population mean (In large samples there is no much diff. b/w sample and population deviations).

Sol Given,

$$n = 900$$

$$\bar{x} = 3.4$$

Scanned with CamScanner

$s = 2.61$ (sample)

$\sigma = 2.61$ (population)

$\mu = 3.25$

i) Null hypothesis $H_0 : \mu = 3.25$

ii) Alternative hypothesis $H_1 : \mu \neq 3.25$

iii) $\alpha = 5\%$.

iv) Test statistics

$$z_{\alpha/2} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \Rightarrow \frac{3.4 - 2.61}{2.61/\sqrt{900}} = 1.75.$$

v) $Z_{tabulated} = 1.96$

conclusion

$z_{cal} < z_{tab} \Rightarrow 1.75 < 1.96$

Accepted H_0.

confidence interval

$$\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

$$= \left(3.4 - 1.75 \times \frac{2.61}{\sqrt{900}}, 3.4 + 1.75 \times \frac{2.61}{\sqrt{900}}\right)$$

$= (\quad)$

9: An ambulance service claims that it takes on the average less than 10 minutes to reach it's destination on emergency calls. A sample of 36 calls has a mean of 11 min and the variance

Scanned with CamScanner

of 16min. Test the claim that 5%. LOS

Sol: Given,

$n = 36$

$\bar{x} = 11$

$s^2 = \sigma^2 = 16$

$\sigma = 4$.

i) Null hypothesis: $H_0 : \mu = 10$

ii) Alternative hypothesis: $H_1 - \mu > 10$.

iii) Level of significance, $\alpha = 5\%$.

iv) Test statistics,

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{11 - 10}{\frac{4}{\sqrt{36}}} = 1.5$$

conclusion,

$z_{tab} = 1.645$

$\therefore z_{cal} < z_{tab}$

$\therefore$ Accepted H_0

Q: The mean lifetime of a sample of 100 light tubes produced by a company is found to be 1560 hours with a population S.D(σ) = 90hrs. Test the significance at 5%. L.O.S that the mean life time of the tubes produced by the company is 1580hrs.

Sol: Given,

$n = 100$

Scanned with CamScanner

$\bar{x} = 1560$

$\sigma = 90$

$\mu = 1580$

i) Null hypothesis : $H_0 = 1580$

ii) Alternative hypothesis, H_1 $\mu < 1580$

iii) L.O.S - $\alpha = 5\%$.

iv) Test hypothesis,

$$z = \frac{1560 - 1580}{\frac{90}{\sqrt{100}}} = -2.22.$$

$z_{tab} = -1.645$

conclusion: $z_{cal} < z_{tab}$

$\therefore$ accepted H_0.

Q: A sample of 400 male students is found to have a mean height 171.38cm. can it be reasonably regarded as a sample from the population with mean height of 171.17cm and S.D of 3.30

Sol: Given,

$n = 400$

$\bar{x} = 171.38$

$\sigma = 3.3$

$\mu = 171.17$

i) Null hypothesis : $H_0 - \mu = 171.17$

ii) Alternative hypothesis : $H_1 - 171.17$

Scanned with CamScanner

iii) L.O.S, $\alpha = 5\%$.

iv) Test significance –

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{171.38 - 171.17}{3.30/\sqrt{400}}$$

$$z = 1.2727$$

Conclusion:

$z_{tab} = 1.96$

$z_{cal} < z_{tab}$

– H_0 is accepted

A sample of 100 ironboxes is drawn from a large no. of balls whose lengths are normally distributed with mean 4 feet and S.D = 0.6 feets. If the sample mean is 4.2 feet. Can the sample be regarded as a truly random sample.

Sol: Given,

$\mu = 4$

$\sigma = 0.6$

$\bar{x} = 4.2$

$n = 100$

i) Null hypothesis, $H_0 - \mu = 4$

ii) Alternative hypothesis, H_1, $\mu \neq 4$.

iii) LOS, $\alpha = 5\%$.

Scanned with CamScanner

iv) Test statistics,

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{4.2 - 4}{0.6/\sqrt{100}}$$

$$z = 3.33.$$

$$z_{cal} = 3.33$$

$$z_{tab} = 1.96.$$

conclusion: $z_{cal} > z_{tab}$

Hence Rejected H_0

Q: The length of life of certain computers is approximately normally distributed with mean 800hrs & S.D 40hrs. If a random sample of 30 computers has an avg. life of 788 hours. Test the Null hypothesis is that $\mu = 800$hrs against the Alternative hypothesis ($\mu \neq 800$hrs, at i) 0.5%. iv) 5%.
ii) 1%. v) 10%.
iii) 4%. vi) 15%.

Sol: Given,

i) 0.5%.

$\alpha = 0.5\%$ (or) 99.5%.

$= 0.995$

$\alpha/2 = 0.995/2 \Rightarrow 0.4975$

$z = 2.81$

ii) 1%.

$\alpha = 1\%$ (or) 99%.

$= 0.99$

Scanned with CamScanner

$\alpha/2 = 0.99/2 \Rightarrow 0.498$

$z = 2.58$

iii) 4%.

$\alpha = 4\%$ or 96%.

$= 0.96$

$\alpha/2 = 0.96/2 = 0.48$

$z = 2.06$.

iv) 5%.

$\alpha = 5\%$ or 95%.

$= 0.95$

$\alpha/2 = 0.95/2 = 0.475$

$z = 1.96$.

TESTING OF HYPOTHESIS.

→ Difference b/w 2 means.

Let n_1, n_2 are the 2 sample sizes having means $\bar{x}$ and $\bar{y}$ drawn from 2 populations with mean μ_1 and μ_2 and of S.D σ_1 & σ_2 respectively.

Now, test the significant difference between the 2 population means.

1) Null hypothesis $H_0 : \mu_1 = \mu_2$

2) Alternative hypothesis, $H_1 : \mu_1 \neq \mu_2$

(i.e $\mu_1 \neq \mu_2$ or $\mu_1 > \mu_2$ or $\mu_1 < \mu_2$)

3) LOS $- \alpha$

Scanned with CamScanner

4) Test statistics,

$$z = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

conclusion.

i) if $z_{cal} < z_{tab} \longrightarrow$ accepted H_0

ii) if $z_{cal} > z_{tab} \longrightarrow$ Rejected H_0

Note

If the 2 samples have been drawn from the 2 populations with common S.D σ then,

$$z = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Q: The mean (~~yield~~) of 2 large samples of sizes 1000 and 2000 are 67.5 inches and 68 inches respectively. Can the samples be regarded as drawn from the same population of S.D 2.5 inches.

Sol: Given,

$n_1 = 1000, n_2 = 2000$

$\bar{x} = 67.5, \bar{y} = 68$

$\sigma = 2.5$

1) Null hypothesis : $H_0 : \mu_1 = \mu_2$

2) Alternative hypothesis : $H_1 : \mu_1 \neq \mu_2$

3) L.O.S $= \alpha = 5\%$.

Scanned with CamScanner

4) Test statistics,

$$z = \frac{(67.5 - 68) - 0}{2.5\sqrt{\frac{1}{1000} + \frac{1}{2000}}}$$

$$z = -5.16$$

$$|z| = 5.16.$$

conclusion;-

$|z|_{tab} = 1.96$

$|z|_{cal} > |z|_{tab}$

Hence reject H_0

Q: Samples of students were drawn from 2 universities and from their weights in kgs, mean and S.D are calculated and shown below. Make a large sample test to test the significance of the difference b/w 2 means.

Mean	S.D	size
55	10	400
57	15	100

Sol: Given,

$n_1 = 400$ $n_2 = 100$

$\bar{x} = 55$ $\bar{y} = 57$

$\sigma_1 = 10$ $\sigma_2 = 15$

1) Null Hypothesis : $H_0 - \mu_1 = \mu_2$

2) Alternative Hypothesis : $H_1 : \mu_1 \neq \mu_2$

Scanned with CamScanner

3) LOS - $\alpha = 5\%$.

4) Test significance:

$$z = \frac{55 - 57}{\sqrt{\frac{10^2}{400} + \frac{15^2}{100}}}$$

$$z = -1.26$$

$$|z|_{cal} = 1.26.$$

$$|z|_{tab} = 1.96.$$

Conclusion: $|z|_{cal} < |z|_{tab}$

Hence Accepted H_0.

Q: The avg. marks scored by 32 boys is 72 with a S.D of 8 while that for 36 girls is 70 with S.D of 6. Does this indicate that the boys perform better than girls at L.O.S 5%.

Sol: Given,

$n_1 = 32 \quad n_2 = 36$

$\bar{x} = 72 \quad \bar{y} = 70$

$\sigma_1 = 8 \quad \sigma_2 = 6$

1) Null Hypothesis - $H_0: \mu_1 = \mu_2$

2) Alternative Hypothesis - $H_1: \mu_1 > \mu_2$

3) L.O.S - $\alpha = 5\%$.

4) Test significance:

$$z = \frac{72 - 70}{\sqrt{\frac{64}{32} + \frac{36}{36}}} = 1.154$$

Scanned with CamScanner

$|z|_{cal} = 1.154$

conclusion: $|z|_{tab} : 1.645$

$|z|_{cal} < |z|_{tab}$

$\therefore$ Accept H_0

Q: A simple sample of the heights of 6400 englishmen has a mean of 67.85 inches and S.D of 2.56 inches. while a simple sample of the heights of 1600 Austrians has a mean of 68.55 inches and S.D of 2.52 inches. Does this data indicate the austrians are on the avg. taller than the englishmen.

Sol: Given,

$n_1 = 6400 \quad n_2 = 1600$

$\bar{x} = 67.85 \quad \bar{y} = 68.55$

$s_1(\sigma_1) = 2.56 \quad s_2(\sigma_2) = 2.52$

1) Null Hypothesis - $H_0: \mu_1 = \mu_2$

2) Alternative Hypothesis - $H_1: (\mu_1 \neq \mu_2) \; \mu_1 < \mu_2$

3) L.O.S - $\alpha = 5\%$

4) Test significance,

$$z = \frac{67.85 - 68.55}{\sqrt{\frac{2.56^2}{6400} + \frac{2.82^2}{1600}}}$$

$z_{cal} = -2.90$

conclusion: $z_{tab} = -1.645$

$z_{cal} < z_{tab}$

$\therefore$ Accepted H_0

Scanned with CamScanner

Q: Suppose that you want to investigate whether on the average man earn more than 20₹ per week than the woman in a certain industry. If sample data shown that 60men earn on the average 292.50 with S.D of 15.6 while 60women earn on average 266.10 with S.D of 18.20 what can we conclude at 1% LOS

Sol: Given,

$\bar{x} = 292.50$ $\quad \bar{y} = 266.10$

$n_1 = 60$ $\quad n_2 = 60$

$s_1(\sigma_1) = 15.6$ $\quad s_2(\sigma_2) = 18.20$.

1) Null Hypothesis - $H_1 : \mu_1 = \mu_2 + 20$

2) Alternative Hypothesis - $\mu_1 \neq \mu_2 + 20$

$\mu_1 > \mu_2 + 20$

3) LOS - $\alpha = 1\%$.

4) Test statistics.

$$z = \frac{(292.50 - 266.10) - (20)}{\sqrt{\frac{15.6^2}{60} + \frac{18.20^2}{60}}}$$

$z_{cal} = 2.068$

$z_{tab} = 2.33$

Conclusion: $z_{cal} < z_{tab}$

$\therefore$ Accept H_0.

Scanned with CamScanner

Test of significance for single proportion.

Suppose a large sample of size "n" (≥ 30) having a proportion 'p' (small letters) which is drawn from the population with proportion P (capital). Test the significance of the population proportion p has a specified value P_0

1) $H_0 : P = P_0$ (i.e $P = Q$)

2) $H_1 : P \neq P_0$ ($P \neq P_0$ or $P > P_0$ or $P < P_0$)

3) LOS $- \alpha$

4) Test statistics : $$z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$$

conclusion:

if $|z|_{cal} < |z|_{tab}$ — Accepted H_0

if $|z|_{cal} > |z|_{tab}$ — Rejected H_0

Q: In a big city 325 men out of 600 men were found to be smokers. Does this information support the conclusion that the majority of men in this city are smokers.

Sol: Given,

$n = 600$

$x = 325$

Scanned with CamScanner

sample proportion for smokers

$$p = \frac{x}{n} = \frac{325}{600} = \frac{13}{24}$$

$$p = 0.5416.$$

population proportion for smokers

$$P = 0.5$$

$$Q = 1 - P = 1 - 0.5$$

$$Q = 0.5$$

1) Null Hypothesis - $H_0 : P = Q = 0.5$

2) Alternative Hypothesis - $H_1 : P > Q$ (or)
$P > 0.5$

3) L.O.S : $\alpha = 5\%$.

4) Test statistics $z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$

$$= \frac{0.54 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{600}}} = 1.95$$

$$z_{cal} = 1.95$$

Conclusion :- $z_{tab} = 1.645$

$$z_{cal} > z_{tab}$$

Hence Rejected H_0

Q: A manufacturer claimed that atleast 95% of the equipment which he supplied to a factory confirm to specification. An examination of a sample of 200 pieces of

Scanned with CamScanner

equipment revealed that 18 were faulty. Test is claimed at 5%. L·O·S

Sol: Given,

$$n = 200$$

Assumed $x = 200 - 18$

items $= 182$

sample proportion for faulty $p = \frac{x}{n}$.

$$= \frac{182}{200} = 0.4$$

population proportion of pieces $P = 0.95$.

$$Q = 1 - P$$
$$= 1 - 0.95$$
$$Q = 0.05$$

1) Null Hypothesis - H_0: $P = 0.95$
2) Alternative Hypothesis - H_1: $P < 0.95$
3) L·O·S : $\alpha = 5\%$.
4) Test significance:

$$z = \frac{0.91 - 0.95}{\sqrt{\frac{0.95 \times 0.05}{200}}}$$

$$z_{cal} = -2.59$$

Conclusion:- $z_{tab} = -1.645$

$$z_{cal} < z_{tab}$$

$\therefore$ Accepted H_0.

Scanned with CamScanner

In a sample of 1000 people in karnataka, 540 are rice eaters and the rest are wheat eaters. Can we assume that both rice and wheat eaters are equally popular in the state at 1%. LOS.

Sol: Given,

$n = 1000$

$x = 540$

$p = \frac{x}{n} = \frac{540}{1000}$

$p = 0.54$

let P = Population proportion for rice eaters $= 0.5$

$Q = 1 - 0.5$

$= 0.5$

1) $H_0 : P = Q = 1/2$

2) $H_1 : P \neq Q$

3) L.O.S : $\alpha = 0.01$

4) Test significance:

$$z = \frac{0.54 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{1000}}} = 2.52$$

Conclusion:

$z_{tab} = 2.58$

$z_{cal} < z_{tab}$

Hence, H_0 Accepted.

Scanned with CamScanner

Q: In a hospital, 480 famales and 520 male babies were born in a week. Do this figures confirm the hypothesis that moles and females are born in equal manner.

Sol; let,

P = population proportion for female babies $= 1/2 = 0.5$

$Q = 1 - P = 1 - 1/2 = 0.5$

total no. of births in a hospital = 1000

p = sample proportion for female babies

$x = 480$

$p = \frac{480}{1000} \Rightarrow 0.48$

1) $H_0 : P = Q = 0.5$

2) $H_1 : P \neq 0.5$ or $P \neq Q$

3) L.O.S - $\alpha = 5\%$

4) Test significance.

$$z = \frac{0.48 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{1000}}} = -1.264$$

Conclusion:

$|z|_{tab} = 1.96$

$|z|_{cal} < |z|_{tab}$

$\therefore$ Accepted H_0

Scanned with CamScanner

Q: 20 people were attacked by a disease and only 18 survived. Will you reject the hypothesis that the survival rate if attacked by this disease 85% in fare are of the hypothesis i.e more at 5% LO.S?

Sol: let,

P = population proportion for survival rate

$= 0.85$

$Q = 1 - P = 1 - 0.85$

$= 0.15$

$x = 18$

$n = 20$

p = sample prop. $= \frac{18}{20} = 0.9$

1) $H_0 : P = 0.85$

2) $H_1 : P > 0.85$

3) L.O.S : $\alpha = 5\%$.

4) Test statistics,

$$z = \frac{0.9 - 0.85}{\sqrt{\frac{0.85 \times 0.15}{20}}} = 0.6262$$

conclusion:

$|z|_{tab} = 1.645$

$|z|_{cal} = 0.6262$

$|z|_{cal} < |z|_{tab}$

$\therefore$ Accepted H_0

Scanned with CamScanner

Q: If 80 patients are treated with an antibiotic 59 got curved. Find a 99% confidence limits for the true population cure if the population of cure is not more.

Sol: Given,

$$n = 80$$

$$x = 59$$

$$P(\text{sample proportion}) = \frac{x}{n} = \frac{59}{80}$$

$$p = 0.7375$$

$$q = 1 - p$$

$$= 1 - 0.7375$$

$$= 0.2625$$

confidence interval,

$$\left(p - z_{\alpha/2}\sqrt{\frac{pq}{n}},\ p + z_{\alpha/2}\sqrt{\frac{pq}{n}}\right)$$

L.O.S = 99%, $\alpha = 1\%$

$$z_{\alpha/2} = 2.58$$

$$= \left(0.7375 - 2.58\sqrt{\frac{0.7375 \times 0.2625}{80}},\right.$$

$$\left. 0.7375 + 2.58\sqrt{\frac{0.7375 \times 0.2625}{80}}\right)$$

$$(0.6, 0.8644)$$

Scanned with CamScanner

Q: A manufacturer claims that only 4% of his products were defective. If random sample for 500 were taken among which 100 were defective. Test the hypothesis at 5% L.O.S

Sol: Given,

$$n = 500$$

$$x = 100$$

$$p = \frac{100}{500} = 0.2$$

$$P = 0.04$$

$$Q = 0.96.$$

1) $H_0 : P = 0.04$

2) $H_1 : P > 0.04$

3) L.O.S $\alpha = 5\%$.

4) Test statistics,

$$z = \frac{0.2 - 0.04}{\sqrt{\frac{0.04 \times 0.96}{500}}} = 18.25$$

$$|z|_{cal} = 18.25$$

Conclusion;

$$|z|_{tab} = 1.645$$

$$|z|_{cal} > |z|_{tab}$$

$\therefore$ Reject H_0

Scanned with CamScanner

Test of significance:

- Difference between 2 proportions.

let P_1, P_2 be the sample proportions in two large samples of sizes n_1 and n_2 drawn from a population with proportions P_1 and P_2. Now, we have to test the

i) the 2 samples have been drawn from the same populations.

ii) is there any significant difference b/w 2 population proportions P_1 and P_2.

1) $H_0 : P_1 = P_2$

2) $H_1 : P_1 \neq P_2$ ($P_1 \neq P_2$ / $P_1 > P_2$ / $P_1 < P_2$)

3) L.O.S: α

4) Test statistics.

i) if population proportions P_1, P_2 known

$$z = \frac{(P_1 - P_2) - (P_1 - P_2)}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}}$$

where, $Q_1 = 1 - P_1$

$Q_2 = 1 - P_2$

ii) If the population proportions. P_1, P_2 are not known then by "Method of Pooling"

Scanned with CamScanner

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$q = 1 - p$$

$$z = \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

conclusion:

$z_{cal} > z_{tab} \Rightarrow$ Reject H_0

$z_{cal} < z_{tab} \Rightarrow$ Accept H_0.

Q: A manufacturer of electronic equipment subjects samples of 2 brands of transistors to a performance test. If 45 of 180 transistors of the first kind and 34 of 120 transistors of the second kind fail the test what can you conclude at 5% LOS about the difference between the corresponding sample proportions.

Sol: Given,

$n_1 = 180 \quad n_2 = 120$

$x_1 = 45 \quad x_2 = 34$

$$p = \frac{x_1 + x_2}{n_1 + n_2} = \frac{45 + 34}{180 + 120}$$

$p = 0.263$

$q = 1 - p$

$= 1 - 0.263$

$= 0.737$

Scanned with CamScanner

1) $H_0 : P_1 = P_2$

2) $H_1 : P_1 \neq P_2$

3) L.O.S - $\alpha = 5\%$.

4) Test statistics,

$$z = \frac{P_1 - P_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{0.25 - 0.28}{\sqrt{0.263 \times 0.737\left(\frac{1}{180} + \frac{1}{120}\right)}}$$

$$z_{cal} = 0.578$$

Conclusion:-

$|z|_{tab} = 1.96$

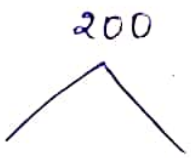
$|z|_{cal} < |z|_{tab}$

$\therefore$ Accept H_0.

Q: On the basis of their total score 200 candidates of a civil service examination are divided into 2 groups, the upper is 30% and remaining 70% considered the first question of the exam. Among the 1st group, 40 had correct answers, whereas in 2nd group 80 had correct answer on the basis of these results can we conclude that the first question is not good at discriminating ability of the type b/w 2 groups.

Scanned with CamScanner

Sol:



$n_1 = 60 \quad n_2 = 140$

$x_1 = 40 \quad x_2 = 80$

$P_1 = 0.566 \quad P_2 = 0.571$

$$P = \frac{x_1 + x_2}{n_1 + n_2} = \frac{40+80}{60+140} = 0.6$$

$$q = 1 - 0.6$$
$$= 0.4$$

1) $H_0 : P_1 = P_2$

2) $H_1 : P_1 \neq P_2$

3) LOS $-\alpha = 5\%$

4) Test statistics :-

$$z = \frac{P_1 - P_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\left(z = \frac{0.25 - 0.28}{\sqrt{0.263 \times 0.737\left(\frac{1}{180} + \frac{1}{120}\right)}}\right)$$

$$= \frac{0.66 - 0.57}{\sqrt{0.6 \times 0.4\left(\frac{1}{60} + \frac{1}{140}\right)}}$$

$z_{cal} = 1.19$

conclusion :- $|z|_{tab} = 1.96$

$|z_{cal}| < |z|_{tab}$

$\therefore$ Accept H_0

Scanned with CamScanner

Q: A cigarette manufacturing company claims that it's brand A of cigarattes outsells it's brand B by 8%. If it is found that 42 out of A sample of 200 smokers prefer brand A and 18 out of 100 smokers prefer brand B. Test whether the 8% difference is

Sol: Given,

$n_1 = 200$ $\quad n_2 = 100$

$x_1 = 42$ $\quad x_2 = 18$

$P_1 = 0.21$ $\quad P_2 = 0.18$

$$P = \frac{x_1 + x_2}{n_1 + n_2} = \frac{42+18}{200+100}$$

$$P = 0.2$$

1) $H_0 : P_1 - P_2 = 8\% = 0.08$

2) $H_1 : P_1 - P_2 \neq 8\%$

3) L.O.S - $\alpha = 5\%$

4) Test statistics, $Z = \dfrac{0.21 - 0.18 - 0.08}{\sqrt{0.2 \times 0.8 \left(\frac{1}{200} + \frac{1}{100}\right)}}$

$$|Z|_{cal} = 1.020$$

Conclusion:-

$|Z|_{cal} < |Z|_{tab}$

$\Rightarrow \therefore$ Accept H_0

Scanned with CamScanner

UNIT-V

SMALL SAMPLES

Small sample.

A random sample of size which is less than 30 ($n < 30$) is known as small sample.

Assumptions for t-Test

The following assumptions are made in t-test.

1) The sample size is less than 30 i.e $n < 30$
2) The population standard Deviation(σ) is not known (σ is unknown)
3) The sample observations are independent

Applications of t-test.

1) To test the specified mean i.e is there any significant difference between a sample mean and the specified value.
2) To test the equality of two means i.e, is there any significant difference between two sample means.
3) To test the significant difference between the means of paired data.

Scanned with CamScanner

Test of significance of single Mean.

A sample of size ($n < 30$) having mean $\bar{x}$ and S.D s which is drawn from population mean μ.

Now, to test, is there any significant difference between sample mean and population mean.

1) Null Hypothesis (H_0): $\mu = \mu_0$ (or) $\bar{x} = \mu$

2) Alternative Hypothesis(H_1): $\mu \neq \mu_0$ (or) $\bar{x} \neq \mu_0$

3) L.O.S :- α

4) Test statistics:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} \qquad \boxed{v = n-1}$$

Conclusion:

i) if $t_{cal} < t_{(n-1)} \rightarrow$ Accept H_0

ii) if $t_{cal} > t_{(n-1)} \rightarrow$ Reject H_0

Confidence Interval:

For true population mean is given by,

i) $\left(\bar{x} - t_{\alpha/2} \frac{s}{\sqrt{n-1}}, \bar{x} + t_{\alpha/2} \frac{s}{\sqrt{n-1}}\right)$

ii) $\left.\begin{aligned} t_1 - \alpha &= -t_\alpha \\ t - \alpha &= -t_\alpha \end{aligned}\right\}$ one-tailed test

Scanned with CamScanner

1) Find the value at

i) $t_{0.05}$ when $n = 16$

$v = n - 1$

$v = 16 - 1 = 15$

$t_{0.05} = t_\alpha$

$\alpha/2 = 0.025$

Now see at 15 in 0.025

$t_{0.05}$ at $v = 15$ is 2.131

(ii) $t_{0.01}$ when $v = 19$.

$t_\alpha = 0.01$

$\alpha/2 = 0.005$

Now see at 19 in 0.005

$t_{0.01}$ at $v = 19$ is 2.861

(iii) $t_{0.995}$ when $v = 7$

$t_{0.995} = t_{1-0.005}$

$= -t_{0.005}$ at $v = 7$ is -3.479.

Since it is one tailed test.

(iv) $t_{-0.01}$ when $n = 9$.

$t_{-0.01} = -t_{0.01}$

at $v = 9$ is -2.896.

Since it is one tailed test.

Scanned with CamScanner

Q: A sample of 26 bulbs gives a mean life of 990hrs with a S.D of 20hrs. The manufacturer claims that the mean life of bulbs is 1000 hours. Is the sample not upto the standard.

Sol: Given,

$$n = 26 (< 30)$$

$$\bar{x} = 990$$

$$s = 20$$

$$\mu = 1000$$

1) $H_0 : \mu = 1000$

2) $H_1 : \mu < 1000$

3) L.O.S $- \alpha = 0.05$

4) Test statistics,

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}}$$

$$= \frac{990 - 1000}{20/\sqrt{26-1}}$$

$$= \frac{-10}{20/\sqrt{25}}$$

$$t = -2.5$$

$$|t|_{cal} = 2.5$$

conclusion:

$$|t|_{tab} = 1.708$$

$$|t|_{cal} > |t|_{tab}$$

$\therefore$ Reject H_0

Scanned with CamScanner

Q: Random sample of size 25 from a normal population as the mean 47.5 and S.D 8.4. Does this information tend to support or refuse the claim that mean of population is 42.5 and also find the confidence interval for the true mean (pop. mean)

Sol: Given,

$n = 25$, $v = 24$

$\bar{x} = 47.5$

$s = 8.4$

$\mu = 42.5$

1) $H_0 : \mu = 42.5$

2) $H_1 : \mu \neq 42.5$

3) L.O.S : $\alpha = 0.05$

4) Test statistics.

$$t = \frac{47.5 - 42.5}{8.4 / \sqrt{25-1}}$$

$$= \frac{47.5 - 42.5}{8.4 / \sqrt{24}}$$

$$t_{cal} = 2.916$$

Conclusion:

$t_{tab} = 2.064$

$\alpha/2 = 0.05/2 = 0.025$

$t_{cal} > t_{tab}$

$\therefore$ Reject H_0

Scanned with CamScanner

Confidence Interval:

$$\left(47.5 - 2.064 \times \frac{8.4}{\sqrt{24}}, 47.5 + 2.064 \times \frac{8.4}{\sqrt{24}}\right)$$

$$= (43.96, 51.03)$$

H_0 does not lies in between.

Q: A machine which produces a mica insulating washer for use in electric devices is set to turn out washers having a thickness of 10mm. A sample of 10 washers has an avg. thickness 9.52mm with a S.D of 0.6mm.

Sol: Given,

$$n = 10, \quad v = 9$$

$$\bar{x} = 9.52$$

$$s = 0.6$$

$$\mu = 10$$

1) $H_0 : \mu = 10$

2) $H_1 : \mu \neq 10$

3) L.O.S - $\alpha = 5\%$.

4) Test statistics,

$$t = \frac{9.52 - 10}{0.6 / \sqrt{10-1}}$$

$$t_{cal} = -2.4$$

$$|t|_{cal} = 2.4$$

Scanned with CamScanner

conclusion:

$t_{tab} = 2.262$

$|t|_{cal} > |t|_{tab}$

Reject H_0

$\alpha = 0.05$

two tailed test

$\alpha/2 = 0.05/2$

$= 0.025$

Q: A random sample of 16 values from a normal population shown a mean of 41.5 inches and sum of the squares of deviation from this mean equals to 135 sq. inches. show that the assumption of a mean of 43.5 inches for the population is not acceptable.

Sol: Given,

$n = 16, v = 15$

$\bar{x} = 41.5$

$\mu = 43.5$

$\Sigma(x_i - \bar{x})^2 = 135$

$$s^2 = \frac{1}{n-1} \Sigma(x_i - \bar{x})^2$$

$$= \frac{1}{16-1} \times 135 = 9$$

$s = 3.$

1) $H_0 : \mu = 43.5$
2) $H_1 : \mu \neq 43.5$
3) L.O.S $- \alpha = 0.05$
4) Test statistics, $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}}$

Scanned with CamScanner

$$= \frac{41.5 - 43.5}{3/\sqrt{16-1}}$$

$$t_{cal} = -2.581$$

$$|t|_{cal} = 2.581$$

conclusion:

$|t|_{tab} = 2.131$ $\quad \alpha = 0.05$

$|t|_{cal} > |t|_{tab}$ $\quad \alpha/2 = 0.025$

$\therefore$ Reject H_0

Q: Prices of shares of a company on the different day in a month were found to be 66, 65, 69, 70, 69, 71, 64, 63 and 68. Discuss whether the mean price of the shares in a month is 65.

Sol: Given,

$\mu = 65, n = 11$

x_i	$X = x_i - \bar{x}$	$(x_i - \bar{x})^2$
66	-1.09	1.1881
65	-2.09	4.3681
69	1.91	3.6481
70	2.91	8.4681
69	1.91	3.6481
71	3.91	15.2821
70	-2.91	8.4681
63	-4.09	16.7281
64	-3.09	9.5481

Scanned with CamScanner

63	-4.09	16.7281
68	0.91	0.8281

$\bar{x} = \frac{\Sigma x_i}{n}$ $\qquad \Sigma(x_i - \bar{x})^2 = 88.903$

$= \frac{1453}{11}$

$= 67.09$

1) $H_0 : \mu = 65$

2) $H_1 : \mu \neq 65$

3) L.O.S - $\alpha = 0.05$

4) Test statistics.

$$t = \frac{67.09 - 65}{2.981/\sqrt{11-1}}$$

$= 2.29.$

$s^2 = \frac{1}{n-1} \times \Sigma(x_i - \bar{x})^2$

$= \frac{1}{11-1} \times 88.903 = 8.8903$

$s = 2.981$

conclusion;

$t_{tab} = 2.228$

$t_{cal} < t_{tab}$

$\therefore$ Accept H_0

Scanned with CamScanner

Q: A random sample of 10 students have following 70, 120, 110, 101, 88, 83, 95, 98, 107, 100.

(i) Do this data support the assumption of a population mean IQ of 100.

(ii) Find a reasonable range in which most of the main IQ values of samples of 10 students lies between.

Sol:

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
70	-27.2	739.84
120	+22.8	519.84
110	12.8	163.84
101	3.8	14.44
88	-9.2	84.64
83	-14.2	201.64
95	-2.2	4.84
98	0.8	0.64
107	9.8	96.04
100	2.8	7.84.

$$\bar{x} = \frac{972}{10}$$

$$= 97.2$$

$$s^2 = \frac{1}{(10-1)} \times 1833.6$$

$$s^2 = 203.73$$

$$s = 10, \nu = 9$$

$$\mu = 100$$

$$\bar{x} = 97.2$$

Scanned with CamScanner

1) $H_0 : \mu = 100$

2) $H_1 : \mu \neq 100$

3) L.O.S $- \alpha = 0.05$

4) Test statistics,

$$t = \frac{97.2 - 100}{14.27 / \sqrt{10-1}} = -0.588$$

$$|t|_{cal} = 0.588$$

conclusion:

$$|t|_{tab} = 2.262$$

$$|t|_{cal} < |t|_{tab}$$

$\therefore$ Accept H_0

confidence Interval

$$\left(97.2 - 2.262 \times \frac{14.27}{\sqrt{9}}, 97.2 + 2.262 \times \frac{14.27}{\sqrt{9}}\right)$$

$$= (86.44, 107.95)$$

H_0 lies in between.

t-Test for: difference of 2 Means.

There are 2 samples of size n_1, n_2 (< 30) having means $\bar{x}$ and $\bar{y}$ with S.D's s_1 and s_2 which are drawn from 2 populations having means μ_1 & μ_2 respectively.

Scanned with CamScanner

To test is there any significant difference between the 2 population means or the 2 samples have come from the same normal population with mean μ.

1) $H_0 : \mu_1 = \mu_2$

2) $H_1 : \mu_1 \neq \mu_2 \;/\; \mu_1 < \mu_2 \;/\; \mu_1 > \mu_2$

3) L.O.S $- \alpha$

4) Test statistics

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where,

$$s^2 = \frac{1}{(n_1 + n_2 - 2)}\left[(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2\right]$$

(or)

$$s^2 = \frac{1}{(n_1 + n_2 - 2)}\left[\Sigma(x_i - \bar{x})^2 + \Sigma(y_i - \bar{y})^2\right]$$

conclusion:

$|t|_{cal} < t_{(n_1+n_2-2)} \longrightarrow$ Accept H_0

$|t|_{cal} > t_{(n_1+n_2-2)} \longrightarrow$ Reject H_0

Note

Confidence Interval for the difference of 2 means (small sample)

Scanned with CamScanner

$$C.I = \left((\bar{x} - \bar{y}) - t_{\alpha/2} \times s \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right.,$$

$$\left. (\bar{x} - \bar{y}) + t_{\alpha/2} \times s \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

Q: The means of 2 random samples of sizes 9 & 7 are 196.42 and 198.82 resp. The sum of the square of deviation from mean are 26.96 and 18.78 resp. can these samples are considered to have been drawn from the same normal population.

Sol: Given,

$n_1 = 9$, $n_2 = 7$

$\bar{x} = 196.42$, $\bar{y} = 198.82$

$\Sigma (x_i - \bar{x})^2 = 26.94$

$\Sigma (y_i - \bar{y})^2 = 18.73$

$$s^2 = \frac{1}{(9+7-2)} [26.94 + 18.73)] = 3.2621$$

$$s = 1.8061$$

1) $H_0 : \mu$ (i.e $\mu_1 = \mu_2$)

2) $H_1 : \mu_1 \neq \mu_2$

3) L.O.S : $\alpha = 0.05$

4) Test statistics,

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Scanned with CamScanner

$$= \frac{196.42 - 198.82}{1.806\sqrt{\frac{1}{9}+\frac{1}{7}}}$$

$t_{cal} = -2.63$

$|t|_{cal} = 2.63$

$\alpha = 0.05$

$\alpha/2 = 0.025$

conclusion:

$|t|_{tab} = 2.145$

$t_{cal} > t_{tab}$

$\therefore$ Reject H_0

$\nu = n_1 + n_2 - 2$

$= 9 + 7 - 2$

$= 14$

Q: 2 Horses A and B were tested acc. to the time (seconds) to run a particular track in the following results. Test whether the 2 horses have the same running capacity.

HorseA	28	30	32	33	33	29	34
HorseB	29	30	30	24	29	29	

Paired Sample t-test.

Paired observations occur in many practical situations like the shooting scores of soldiers before and after training or the performance of the students in competitive exams before and after coaching etc.

If $(x_1, y_1)(x_2, y_2) \ldots$ be the paired observations for the given data then,

Scanned with CamScanner

1) $H_0 : \mu_1 = \mu_2$ (i.e $\mu = 0$)

[There is no significant difference between the means of two situations].

2) $H_1 : \mu_1 < \mu_2$

3) L.O.S $- \alpha$

4) Test statistics,

$$t = \frac{\bar{d} - \mu}{s/\sqrt{n-1}} \quad ; \quad d_i = \bar{x}_i - y_i \quad [\because \mu = 0]$$

conclusion:

$t_{cal} < t_{tab} \longrightarrow$ accept H_0

$t_{cal} > t_{tab} \longrightarrow$ reject H_0

Q: Scores obtained in a shooting competition by 10 soldiers before and after the intensive training are given

Before	67	24	57	55	63	54	56	68	33	43
After	70	38	58	58	56	67	68	75	42	38

Test whether the intensive training is useful or not

Sol: Given,

$n = 10$

Scanned with CamScanner

Before	After	$d_i = x_i - y_i$	$d_i - \bar{d}$	$(d_i - \bar{d})^2$
67	70	−3	2	4
24	38	−14	−9	81
57	58	−1	4	16
55	58	−3	2	4
63	56	7	12	144
54	67	−13	−8	64
56	68	−12	−7	49
68	75	−7	−2	4
33	42	−9	−4	16
43	38	5	10	100
				482

$$\bar{d} = \frac{\Sigma d_i}{n}$$

$$= -5$$

1) $H_0 : \mu = 0$

2) $H_1 : \mu < 0$

3) L.O.S $\alpha = 0.05$

4) Test statistics.

$$t = \frac{5 - 0}{7.31 / \sqrt{10 - 1}}$$

$$t_{cal} = -2.051 \qquad |t_{cal}| = 2.051$$

conclusion:

[t_α] $t_{tab} = 1.833$ at $\nu = 9$

Scanned with CamScanner

$t_{cal} > t_{tab}$

Reject H_0

Training is not useful.

Q: The blood pressure of 5 persons before and after intake of a certain drug are given below.

Before	110	120	125	132	125
After	120	118	125	136	121

Test whether there is significant change in B.P at 1% LOS.

Before	After	$d_i = x_i - y_i$	$d_i - \bar{d}$	$(d_i - \bar{d})^2$
110	120	-10	-8.4	70.56
120	118	2	3.6	12.96
125	125	0	1.6	2.56
132	136	-4	-2.4	5.76
125	121	4	5.6	31.36
				$\Sigma(d_i - \bar{d}) =$ 123.2

$\bar{d} = \frac{-8}{5}$

$= -1.6$

1) $H_0 : \mu_1 = 0$

2) $H_1 : \mu_1 < 0$

3) L.O.S $- \alpha = 0.01$

$s^2 = \frac{1}{5-1} \times 123.2$

$= 30.8$

$s = 5.549$

Scanned with CamScanner

4) Test statistics-

$$t = \frac{-1.6 - 0}{5.549/\sqrt{4}}$$

$$t_{cal} = -0.576$$

$$|t|_{cal} = 0.576$$

conclusion:

$t_{0.01}$ at $v = 4$ is 3.747

$$|t|_{tab} = 3.747$$

$$t_{cal} < t_{tab}$$

Accept H_0

Q: Memory capacity of 10 students were tested before and after the training. state whether the training was effective or not from the following scores.

Before	After.
12	15
14	16
11	18
8	10
7	5
10	12
3	10
0	2
5	3
6	8

Scanned with CamScanner

t-test for Difference of 2 Means.

Q: Measuring specimens of nylon yarn taken from 2 machines it was found that 8 specimens from 1st machine had a mean of 9.67 with S.D of 1.81 while 10 specimens from second machine had a mean of 7.43 with S.D of 1.48. Assuming that the proportions are normal, Test the hypothesis, at 5% L.O.S

Sol: Given,

$n_1 = 8 \quad n_2 = 10$

$\bar{x} = 9.67 \quad \bar{y} = 7.43$

$s_1 = 1.81 \quad s_2 = 1.48$

1) $H_0 : \mu_1 - \mu_2 = 1.5$

2) $H_1 : \mu_1 - \mu_2 > 1.5$

3) L.O.S $- \alpha = 5\%$

4) Test statistics.

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s^2 = \frac{1}{8+10-2}\left[(8-1)(1.81)^2 + (10-1)(1.48)^2\right]$$

$$s^2 = 2.66$$

$$s = 1.63$$

$$= \frac{9.67 - 7.43 - 1.5}{1.63\sqrt{\frac{1}{8} + \frac{1}{10}}}$$

$$t_{cal} = 0.95$$

Scanned with CamScanner

conclusion:-

$t_{tab} = 1.746$

$t_{cal} < t_{tab}$

Accept H_0

at $v = n_1 + n_2 - 2$

$= 8 + 10 - 2$

$= 16$

Q: The IQ's of 16 students from one area of a city shown a mean of 107 with a S.D of 10 while IQ's of 14 students from another area of a city shown a mean of 112 with a S.D of 8. Is there any significant difference between the IQ's of the 2 groups at 5% L.O.S?

F-test (snedecor's F-test) - (variances)

F-test is used for the equality of population variances.

Test of significance of 2 population variances. Suppose 2 samples of sizes n_1 and n_2 (< 30) having variances s_1^2 and s_2^2 which are drawn from 2 normal populations having variances σ_1^2 and σ_2^2. To test there is any significant difference between the population variances or the 2 samples have been drawn from the normal population with same variance.

Scanned with CamScanner

n_1 | n_2

s_1^2 | s_2^2

σ_1^2 | σ_2^2

$H_0: \sigma_1^2 = \sigma_2^2 = \sigma^2$

$H_1: \sigma_1^2 \neq \sigma_2^2$

Test statistics,

$$F = \frac{\text{Greater variance}}{\text{smaller variance}}$$

i.e $F = \frac{s_1^2}{s_2^2}$ if $s_1^2 > s_2^2$

or $F = \frac{s_2^2}{s_1^2}$ if $s_2^2 > s_1^2$

conclusion:

$F_{cal} < F_{tab} \longrightarrow$ accept H_0

$F_{cal} > F_{tab} \longrightarrow$ Reject H_0

$F_{0.05}(v_1, v_2)$ $F_{0.95}$

$F_{0.01}(v_1, v_2)$ $F_{0.9}$

$$F_{1-\alpha}(v_1, v_2) = \frac{1}{F_{\alpha}(v_2, v_1)}$$

Scanned with CamScanner

Find,

i) $F_{0.05}$ with $v_1 = 7, v_2 = 15$

$= 2.7$

ii) $F_{0.095}$ with $v_1 = 19$ and $v_2 = 20$

$$F_{0.95}(19,24) = F_{1-0.05}(19,24)$$

$$= \frac{1}{F_{0.05}(24,19)}$$

$$= \frac{1}{2.11} = 0.473.$$

(iii) $F_{0.01}$ with $n_1 = 25, n_2 = 20$

$v_1 = 24, v_2 = 19$ is 2.92

(iv) $F_{0.99}$ with $v_1 = 28$ & $v_2 = 12$

$$F_{0.99}(28,12) = F_{1-0.01}(28,12)$$

$$= \frac{1}{F_{0.01}(12,28)}$$

$$= \frac{1}{2.90} = 0.34$$

Q: In one sample of 8 observations from a normal observations the sum of the squares of the deviations from the sample mean is 84.4 and in another sample of 10 observations it was 102.6. Test at 5% L.O.S whether the populations have the same variance.

Scanned with CamScanner

$n_1 = 8$ $\quad n_2 = 10$

$\Sigma(x_i - \bar{x})^2 = 84.4$ $\quad \Sigma(y_i - \bar{y})^2 = 102.6$

$S_1^2 = \frac{1}{8-1} \times 84.4$ $\quad S_2^2 = \frac{1}{10-1} \times 102.6$

$S_1^2 = 12.057$ $\quad S_2^2 = 11.4$

1) $H_0 : \sigma_1^2 = \sigma_2^2$

2) $H_1 : \sigma_1^2 \neq \sigma_2^2$

3) L.O.S - $\alpha = 0.05$

4) Test statistics :-

$$F = \frac{S_1^2}{S_2^2} = \frac{12.05}{11.44} \quad (S_1^2 > S_2^2)$$

$$= 1.057$$

conclusion:

$$F_{tab} = F_{0.05}(7,9) = 3.29$$

$\therefore F_{cal} < F_{tab}$

Accept H_0

Q: The time taken by workers in performing a job by method 1 and method 2 is given below.

Method 1	20	16	26	27	23	22	
Method 2	27	33	42	35	32	34	38

Scanned with CamScanner

Q: Test whether a nicotin contains in mgs in 2 samples of tobacco were found to be as follows.

sample A	24	27	26	21	25	
sample B	27	30	28	31	22	36.

can it be said that the 2 samples have come from same normal population.

Sol:

Sample A	sample B	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	$y_i-\bar{y}$	$(y_i-\bar{y})^2$
24	27	−0.6	0.36	−2	4
27	30	2.4	5.76	1	1
26	28	1.4	1.96	−1	1
21	31	−3.6	12.96	2	4
25	22	0.4	0.16	−7	49
	36			7	49
		21.2			108

$$\bar{x} = \frac{\Sigma x}{n} = \frac{123}{5} = 24.64$$

$$\bar{y} = \frac{\Sigma y}{n} = 29.$$

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

L.O.S $-\alpha = 0.05$

$$s^2 = \frac{1}{n_1+n_2-2}\left(\Sigma(x_i-\bar{x})^2 + \Sigma(y_i-\bar{y})^2\right)$$

$$= \frac{1}{5+6-2}(21.2+108)$$

$$= 14.355$$

$$S = 3.788$$

Scanned with CamScanner

Test statistics:

$$t_{cal} = \frac{24.6 - 29}{3.788\sqrt{\frac{1}{5}+\frac{1}{6}}}$$

$$= -1.912$$

$$|t_{cal}| = 1.912$$

$t_{tab} = 2.262$ at $v = 9$.

$\alpha/2 = 0.05/2 = 0.025$

$t_{cal} < t_{tab}$

Accept H_0

χ^2-test (chi square test)

χ^2 distribution is a continuous probability distribution and it is used in both large and small samples. The χ^2-test is mainly used in,

i) to test the goodness of fit
ii) to test the independence of attributes
iii) to test if the population variance has a specified value.

χ^2-test for Goodness of fit

Suppose $O_1, O_2, O_3 \cdots O_i \cdots$ be the observed frequencies and $E_1, E_2 \cdots E_i \cdots$ be the corresponding expected frequencies of some experiment then, the χ^2-test

Scanned with CamScanner

is given by,

$$\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$$

H_0: There is no significant difference between the observed frequencies and expected frequencies.

H_1: There is some significant difference between the observed and expected frequencies.

L·O·S - α

Test statistics: $\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$

conclusion:

$\chi^2_{cal} < \chi^2_{tab} \rightarrow$ Accept H_0

$\chi^2_{cal} > \chi^2_{tab} \rightarrow$ Reject H_0

Q: Fit a Poisson distribution to the following data and for it's goodness of fit at 5% L·O·S?

x	0	1	2	3	4
f	419	352	154	156	90

Sol: $\lambda = \frac{\Sigma x_i f_i}{\Sigma f_i}$

$= \frac{1788}{13.42}$

$\lambda = 1.33$

Scanned with CamScanner