UNIT-V

STOCHASTIC PROCESS

Stochastic process (or) Random process

The family of all the random variables at particular time 't' is known as stochastic process (or) random process.

It is denoted with $X(t_n)$ for $n = 1, 2, 3, \ldots$

$t_1\ t_2\ t_3 \ldots$ are called states of the stochastic process.

Ex:- Poisson distribution is a stochastic process with infinite no. of states.

$$P_n(t) = \frac{e^{-\lambda t}(\lambda t)^n}{n!} \quad \forall\, n = 0, 1, 2, 3, \ldots$$

Markov property: A property is said to be Markov property if the occurance of future state depends on immediately preceeding state.

ie; $\{X(t_n) = x_n / X(t_{n-1}) = x_{n-1}\}$

Ex:- The probability of being dry tomarrow is 0.8 if today is dry, but it is only 0.6 if it rains today.

Markov process (or) Markov Chain:- A stochastic process is said to be Markov process (or) Markov chain if it ~~s~~ satisfies Markov property.

ie; $\{X(t_n) = x_n / X(t_{n-1}) = x_{n-1}\}$

Ex:- A game of ~~s~~ snakes and ladders whose moves are determined entirely by dice is a markov chain.

Transition probability:- The probability of future state is depends on immediately preceding state is known as transition probability.

$$P_{x_{n-1} \to x_n} = P\{X(t_n) = x_n / X(t_{n-1}) = x_{n-1}\}$$

Transition probability matrix:- The transition probabilities can be arranged in matrix form such a matrix is called transition probability matrix.

It is denoted by $P = \begin{bmatrix} P_{11} & P_{12} & P_{13} & \cdots & P_{1n} \\ P_{21} & P_{22} & P_{23} & \cdots & P_{2n} \\ \vdots & & & & \\ P_{n1} & P_{n2} & P_{n3} & \cdots & P_{nn} \end{bmatrix}$

In General the matrix is said to be transition probability matrix if it satisfies the following properties.

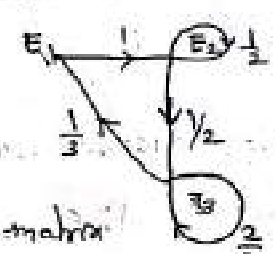
i) it is a square matrix with non negative elements

ii) sum of each row is equal to 1.

Transition Diagram

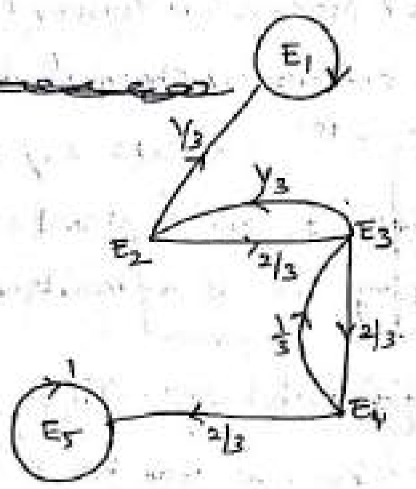
The diagram of transition probability matrix is known as transition diagram.

Ex:- ① $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{2}{3} \end{bmatrix}$ $\begin{array}{c|ccc} & E_1 & E_2 & E_3 \\ \hline E_1 & 0 & 1 & 0 \\ E_2 & 0 & \frac{1}{2} & \frac{1}{2} \\ E_3 & \frac{1}{3} & 0 & \frac{2}{3} \end{array}$

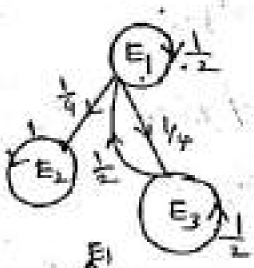


A zero element in the above transition probability matrix indicates that the transition is not possible

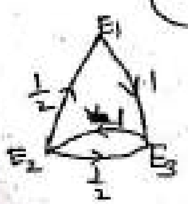
Ex:- ② $P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ $\begin{array}{c|ccccc} & E_1 & E_2 & E_3 & E_4 & E_5 \\ \hline E_1 & 1 & 0 & 0 & 0 & 0 \\ E_2 & \frac{1}{3} & 0 & \frac{2}{3} & 0 & 0 \\ E_3 & 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ E_4 & 0 & 0 & \frac{1}{3} & 0 & \frac{2}{3} \\ E_5 & 0 & 0 & 0 & 0 & 1 \end{array}$



Ex:- ③ $A = \begin{array}{c|ccc} & E_1 & E_2 & E_3 \\ \hline E_1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ E_2 & 0 & 1 & 0 \\ E_3 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}$



Ex:- ④ $A = \begin{array}{c|ccc} & E_1 & E_2 & E_3 \\ \hline E_1 & 0 & 0 & 1 \\ E_2 & \frac{1}{2} & 0 & \frac{1}{2} \\ E_3 & 0 & 1 & 0 \end{array}$



Types of Markov Chain

1) Ergodic Markov Chain
2) Regular Markov Chain

1) **Ergodic Markov Chain:—** A markov chain is said to be Ergodic if it has a property i.e. possible to pass from all the present states to all the future states in finite no. of steps.

2) **Regular Markov Chain:—** A markov chain is said to be regular if some power of transition matrix P becomes non-zero matrix.

Note:— All the regular chains are Ergodic chains.

Types of matrices

1) Stochastic matrix 2) Regular stochastic matrix.

1) **Stochastic matrix:—** A matrix is said to be stochastic matrix if it satisfies following properties.
   i) it is a square matrix with non negative elements
   ii) the sum of each row is 1.

2) **Regular stochastic matrix:—** A matrix is said to be regular stochastic matrix if some powers of P becomes non-zero matrix.

① Which of the following matrices are stochastic

i) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

Which is not stochastic because it is not a square matrix.

ii) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Which is stochastic because it is a square matrix and each row's sum is equal to 1.

iii) $\begin{bmatrix} 0 & 1 \\ \frac{1}{3} & \frac{1}{4} \end{bmatrix}$

Which is not stochastic matrix because sum of each row is not equal to 1

iv) $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

Which is a stochastic because it is a square matrix and each row's sum is equal to 1

v) $\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$

which is not stochastic because there is a negative element.

③ Which of the following are regular stochastic matrices

i) $A = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$

It is a stochastic matrix, but

$A^2 = \begin{pmatrix} \frac{3}{8} & \frac{3}{8} & \frac{1}{4} \\ 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{8} & \frac{3}{8} \end{pmatrix}$ $A^3 = \begin{pmatrix} \frac{5}{16} & \frac{15}{32} & \frac{7}{32} \\ 0 & 1 & 0 \\ \frac{7}{16} & \frac{1}{4} & \frac{5}{16} \end{pmatrix}$

Here zeros are repeating in $A^2, A^3 \ldots$

$\therefore$ It is not a regular stochastic matrix.

ii) $B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}$

It is a stochastic matrix

$B^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{3}{4} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ $B^3 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{5}{8} & \frac{3}{8} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{pmatrix}$ $B^4 = \begin{pmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ 0 & \frac{11}{16} & \frac{5}{16} \\ 0 & \frac{5}{8} & \frac{3}{8} \end{pmatrix}$ $B^5 = \begin{pmatrix} 0 & \frac{5}{8} & \frac{3}{8} \\ 0 & \frac{21}{32} & \frac{11}{32} \\ 0 & \frac{11}{16} & \frac{5}{16} \end{pmatrix}$

Here zeros are repeating in $B^2, B^3, B^4, B^5, \ldots$

$\therefore$ It is not a regular stochastic matrix

iii) $C = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}$

It is a stochastic, $C^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$ $C^3 = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ $C^4 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$

$C^5 = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{8} & \frac{1}{2} & \frac{3}{8} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$ $\therefore$ Here ~~powers~~ of C^5 ~~are~~ is not zero matrix : $\therefore$ It is regular

iv) $D = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$ $D^2 = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} \\ 0 & 1 \end{pmatrix}$ $D^3 = \begin{pmatrix} \frac{1}{8} & \frac{7}{8} \\ 0 & 1 \end{pmatrix}$ $\therefore$ It is not regular

v) $E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $E^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $E^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ $E^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ not regular.

vi) $F = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$ $F^2 = \begin{pmatrix} \frac{3}{8} & \frac{1}{2} & \frac{1}{8} \\ 0 & 1 & 0 \\ \frac{1}{4} & \frac{5}{8} & \frac{1}{8} \end{pmatrix}$ $F^3 = \begin{pmatrix} \frac{1}{4} & \frac{21}{32} & \frac{3}{32} \\ 0 & 1 & 0 \\ \frac{3}{16} & \frac{3}{4} & \frac{1}{16} \end{pmatrix}$ not regular

vii) $G = \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}$ $G^2 = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{8} & \frac{5}{16} & \frac{9}{16} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$ $G^3 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{5}{32} & \frac{41}{64} & \frac{13}{64} \\ \frac{1}{8} & \frac{5}{16} & \frac{9}{16} \end{pmatrix}$ regular.

③ Check whether the following ~~matrix of~~ Markov chain is regular and ergodic

i) $P = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

$P^2 = \begin{pmatrix} \frac{3}{8} & \frac{1}{8} & \frac{1}{8} & \frac{3}{8} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$

$\therefore$ It is non zero matrix so P is regular chain

ii) $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$

$P^2 = \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ $P^3 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$ $P^4 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$

$P^5 = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{8} & \frac{3}{8} & \frac{1}{2} \end{pmatrix}$

$\therefore$ It is nonzero matrix So P is regular chain

Check whether the chain is Regular and Ergodic

$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

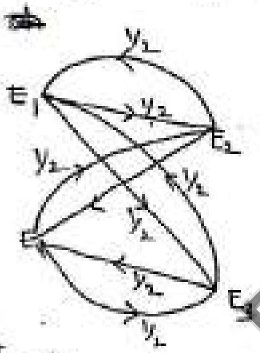
i:- $P^2 = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}$ $P^3 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$ $P^4 = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}$

$\therefore$ Zero elements are repeated, so it is not regular.

Consider

$$P = \begin{array}{c|cccc} & E_1 & E_2 & E_3 & E_4 \\ E_1 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ E_2 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ E_3 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ E_4 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}$$

Sol:-



In this matrix it is possible from all the present states to future states in finite no. of steps.

$\therefore$ It is Ergodic.

State to state

$E_1 \to E_2 \to E_1$
(or)
$E_1 \to E_3 \to E_1$
$E_1 \to E_2$
$E_1 \to E_3$
$E_1 \to E_3 \to E_4$
(or)
$E_1 \to E_2 \to E_4$
$E_2 \to E_1$
$E_2 \to E_1 \to E_2$
(or)
$E_2 \to E_4 \to E_2$
$E_2 \to E_1 \to E_3$
$E_2 \to E_4$
$E_3 \to E_1$
$E_3 \to E_1 \to E_2$
(or)
$E_3 \to E_4 \to E_3$
(or)
$E_3 \to E_1 \to E_3$
$E_3 \to E_4$
$E_4 \to E_2 \to E_1$
(or)
$E_4 \to E_3 \to E_1$
$E_4 \to E_2$
$E_4 \to E_3$
$E_4 \to E_2 \to E_4$
$E_4 \to E_3 \to E_4$

iv) $D = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$ $D^2 = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} \\ 0 & 1 \end{pmatrix}$ $D^3 = \begin{pmatrix} & \\ 0 & 1 \end{pmatrix}$

(5) Check whether the following ~~matrix~~ markov chain is regular and ergodic.

Sol: $P = \begin{bmatrix} 0 & x & x & 0 \\ x & 0 & 0 & x \\ x & 0 & 0 & x \\ 0 & x & x & 0 \end{bmatrix}$

Sol: $P^2 = \begin{bmatrix} 2x^2 & 0 & 0 & 2x^2 \\ 0 & 2x^2 & 2x^2 & 0 \\ 0 & 2x^2 & 2x^2 & 0 \\ 2x^2 & 0 & 0 & 2x^2 \end{bmatrix}$ $P^3 = \begin{bmatrix} 0 & 4x^3 & 4x^3 & 0 \\ 4x^3 & 0 & 0 & 4x^3 \\ 4x^3 & 0 & 0 & 4x^3 \\ 0 & 4x^3 & 4x^3 & 0 \end{bmatrix}$

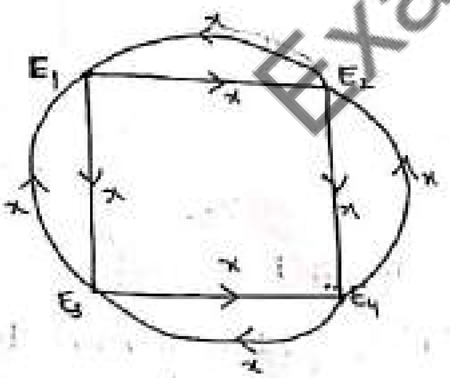
$P^4 = \begin{bmatrix} 8x^4 & 0 & 0 & 8x^4 \\ 0 & 8x^4 & 8x^4 & 0 \\ 0 & 8x^4 & 8x^4 & 0 \\ 8x^4 & 0 & 0 & 8x^4 \end{bmatrix}$

$\therefore$ It is not regular

Consider

$$P = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \\ \end{array}$$

$$P = \begin{array}{c|cccc} & E_1 & E_2 & E_3 & E_4 \\ \hline E_1 & 0 & x & x & 0 \\ E_2 & x & 0 & 0 & x \\ E_3 & x & 0 & 0 & x \\ E_4 & 0 & x & x & 0 \end{array}$$



In this matrix it is possible from all the present states to future states in finite no of steps.

$\therefore$ It is Ergodic.

state - to - state

$E_1 \to E_2 \to E_1$ $E_1 \to E_3 \to E_1$

$E_1 \to E_2$ $E_1 \to E_3$

$E_1 \to E_3 \to E_4$, $E_1 \to E_2 \to E_4$

$E_2 \to E_1$

$E_2 \to E_1 \to E_2$ $E_2 \to E_4 \to E_2$

$E_2 \to E_4 \to E_3$ $E_2 \to E_1 \to E_3$

$E_2 \to E_4$

$E_3 \to E_1$

$E_3 \to E_1 \to E_2$ $E_3 \to E_4 \to E_2$

$E_3 \to E_4 \to E_3$ $E_3 \to E_1 \to E_3$

$E_3 \to E_4$

$E_4 \to E_2 \to E_1$ $E_4 \to E_3 \to E_1$

$E_4 \to E_2$

$E_4 \to E_3$

$E_4 \to E_3 \to E_4$ $E_4 \to E_2 \to E_4$

⑥ Check the following matrix is regular and Ergodic

$$A = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

Sol:-

$$A^2 = \begin{bmatrix} \frac{1}{9} & \frac{13}{36} & \frac{17}{72} & \frac{11}{24} \\ \frac{1}{16} & \frac{1}{4} & \frac{1}{3} & \frac{17}{48} \\ \frac{5}{24} & \frac{1}{12} & \frac{1}{3} & \frac{3}{8} \\ \frac{1}{12} & 0 & \frac{7}{18} & \frac{19}{36} \end{bmatrix}$$

A^3 is nonzero, so A is Ergodic matrix.

Classification of states

Consider a 4×4 matrix

$$P = \begin{array}{c|c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \\ P_{41} & P_{42} & P_{43} & P_{44} \end{bmatrix} \end{array}$$

1) **Absorbing state**:- If $P_{ii} = 1$ then i is said to be absorbing state.

Ex:-

$$\begin{bmatrix} \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Here absorbing state is 3, because $P_{33} = 1$

2) **transient state**:- If $P_{ii} < 1$ then state 'i' is said to be transient state

From the above example, the transient states are 1, 2, 4, 5

3) **Return state**:- If $P_{ii}^{(n)} > 0$ for some n then 'i' is called return state.

From the above example, $P_{00} > 0$, $P_{11} > 0$, $P_{22} > 0$, $P_{33} > 0$ for $P^{(1)}$ But $P_{44}^{(n)} \neq 0$.

Here return states are 1, 2, 3, 4

(iv) $D = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$ $D^2 = \begin{bmatrix} \cdot & \cdot \\ 0 & 1 \end{bmatrix}$ $D = \begin{bmatrix} \cdot & \cdot \\ 0 & 1 \end{bmatrix}$

(4) Irreducible state of matrix :- If $P_{ij}^{(n)} > 0$ for some n then it is irreducible matrix otherwise reducible matrix.

Ex:- The transition probability matrix of a markov chain is given by

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0.3 & 0.7 & 0 \\ 0.1 & 0.4 & 0.5 \\ 0 & 0.2 & 0.8 \end{bmatrix} \end{array}.$$

Here $P_{11}^{(1)} > 0$, $P_{12}^{(1)} > 0$, $P_{21}^{(1)} > 0$ $P_{22}^{(1)} > 0$ $P_{23}^{(1)} > 0$, $P_{32}^{(1)} > 0$, $P_{33}^{(1)} > 0$

But $P_{13}^{(1)} = 0$, $P_{31}^{(1)} > 0$

$$P^2 = \begin{bmatrix} 0.16 & 0.49 & 0.35 \\ 0.07 & 0.35 & 0.6 \\ 0.02 & 0.24 & 0.74 \end{bmatrix}$$

Here $P_{13}^{(2)} > 0$ & $P_{31}^{(2)} > 0$ so it is irreducible

(1) Consider the markov chain with transition probability matrix.

$$P = \begin{bmatrix} 0.4 & 0.6 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 \\ 0.2 & 0.4 & 0.1 & 0.3 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$ Is it irreducible ?

Sol:- Here $P_{11}^{(1)} > 0$, $P_{12}^{(1)} > 0$, $P_{21}^{(1)} > 0$, $P_{22}^{(1)} > 0$, $P_{31}^{(1)} > 0$, $P_{32}^{(1)} > 0$, $P_{33}^{(1)} > 0$, $P_{34}^{(1)} > 0$

$P_{44}^{(1)} > 0$

$$P^{(2)} = \begin{bmatrix} 0.34 & 0.66 & 0 & 0 \\ 0.21 & 0.67 & 0 & 0 \\ 0.22 & 0.44 & 0.01 & 0.03 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Zero's are repeated same places, so P is reducible.

Periodic and aperiodic states

The periodic of return state is defined as the G.C.D. of all 'n' such that $P_{ij}^{(n)} > 0$; $d_i = GCD\{n, P_{ii}^{(n)} > 0\}$

If $d_i > 1$ then state 'i' is called periodic state.

If $d_i = 1$ then state 'i' is called aperiodic state.

① A transition probability matrix is $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$, Check whether the states are periodic (or) aperiodic.

Sol:- Given $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$ $\quad P^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \end{bmatrix}$

$P^3 = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$ $\quad P^4 = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.25 & 0.25 & 0.5 \\ 0.25 & 0.5 & 0.25 \end{bmatrix}$

$P^5 = \begin{bmatrix} 0.25 & 0.25 & 0.5 \\ 0.25 & 0.5 & 0.25 \\ 0.125 & 0.375 & 0.5 \end{bmatrix}$

$P_{11}^{(3)} > 0$, $P_{11}^{(5)} > 0$

$d_i = G.C.D(3, 5) = 1$

~~state is aperiodic~~ "1" is aperiodic state,

$P_{22}^{(2)} > 0$, $P_{21}^{(3)} > 0$, $P_{22}^{(4)} > 0$, $P_{22}^{(5)} > 0$

$d_i = G.C.D(2, 3, 4, 5) = 1$

"2" ~~state~~ is aperiodic state

$P_{33}^{(2)} > 0$, $P_{33}^{(3)} > 0$, $P_{33}^{(4)} > 0$, $P_{33}^{(5)} > 0$

$d_i = G.C.D(2, 3, 4, 5) = 1$

"3" ~~state~~ is aperiodic state.

iv) $D = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$

② 3 boys A, B, C are throwing a ball to each other, A always throws the ball to B and B always throws the ball to C but C just as likely to B as to A. Show that the process is Markov. Find the transition matrix and classify the states. Check the matrix is ergodic or not?

Sol:- Given

$$P = \begin{array}{c c} & \begin{array}{ccc} A & B & C \end{array} \\ \begin{array}{c} A \\ B \\ C \end{array} & \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \end{array}$$

The future state are always depends on present state.

$\therefore$ The chain is called markov chain.

Here A, B, C states are not absorbing states ($P_{ii} \neq 1$)

A, B, C states are transient states ($P_{ii} < 1$)

$$P^2 = \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad P_{33}^{(2)} > 0, \; P_{22}^{(2)} > 0 \quad B \,\&\, C \text{ are return states}$$

$$P^3 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix} \quad P_{11}^{(3)} > 0 \quad A \text{ is return state.}$$

$P_{ij}^{(n)} > 0$ i.e. $P_{12}^{(1)} > 0$, $P_{23}^{(1)} > 0$, $P_{31}^{(1)} > 0$, $P_{32}^{(1)} > 0$

$P_{13}^{(2)} > 0$, $P_{21}^{(2)} > 0$, $P_{22}^{(2)} > 0$, $P_{33}^{(2)} > 0$, $P_{11}^{(3)} > 0$

$\therefore$ P is an Irreducible matrix

$$P^{(4)} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix} \qquad P^5 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{8} & \frac{3}{8} & \frac{1}{2} \end{pmatrix}$$

It is a regular matrix & P is ergodic matrix.

$P_{11}^{(3)} > 0$, $P_{11}^{(5)} > 0$

$d_i = G.C.D\{3, 5\} = 1$

$\therefore$ A is aperiodic state.

$P_{22}^{(2)} > 0 \quad P_{22}^{(3)} > 0, \; P_{22}^{(4)} > 0, \; P_{22}^{(5)} > 0$

$d_i = G.C.D\{2,3,4,5\}$

$d_i = 1$

$\therefore$ B is aperiodic state.

$P_{33}^{(2)} > 0, \; P_{33}^{(3)} > 0, \; P_{33}^{(4)} > 0, \; P_{33}^{(5)} > 0$

$d_i = G.C.D\{2,3,4,5\}$

$d_i = 1$

$\therefore$ C is aperiodic state.

③ Find the periodic and aperiodic states in the following transition probability matrices

i) $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

Sol:- $A^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad A^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad A^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad A^5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

$P_{11}^{(2)} > 0, \; P_{11}^{(4)} > 0 \quad d_i = G.C.D\{2,4\} = 2$

$\therefore$ 1 is periodic state.

$P_{22}^{(2)} > 0, \; P_{22}^{(4)} > 0 \quad d_i = G.C.D\{2,4\} = 2$

$\therefore$ 2 is periodic state.

ii) $P = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

$P^2 = \begin{bmatrix} 0.4375 & 0.5625 \\ 0.375 & 0.625 \end{bmatrix} \quad P^3 = \begin{bmatrix} 0.3906 & 0.6093 \\ 0.4062 & 0.5937 \end{bmatrix}$

$P^4 = \begin{bmatrix} 0.4023 & 0.5976 \\ 0.3984 & 0.6015 \end{bmatrix} \quad P^5 = \begin{bmatrix} 0.3994 & 0.6005 \\ 0.4003 & 0.5996 \end{bmatrix}$

$P_{11}^{(2)} > 0, \; P_{11}^{(3)} > 0, \; P_{11}^{(4)} > 0, \; P_{11}^{(5)} > 0 \quad d_i = G.C.D\{2,3,4,5\} = 1$

$P_{22}^{(2)} > 0 \quad P_{22}^{(3)} > 0 \quad P_{22}^{(4)} > 0 \quad P_{22}^{(5)} > 0 \quad d_i = G.C.D\{2,3,4,5\} = 1$

$\therefore$ 1, 2, states are aperiodic states

Gambler's ruin problem

Suppose a Gambler wins or looses a unit (either it is rupees or dollar) with probability p and q respectively.

Let his initial capital is 'z' and opponent's initial capital is '(a−z)'

Then the combined capital is $z + a - z = a$

The Game continuous either the Gambler's capital reduced to zero or increased to a (ie; 1 of the two players will be ruined), we can find out-

i) probability of Gambler's ruined (q_z)

ii) probability of opponent's gambler ruined (p_z)

iii) probability of duration of the game in two cases, biased and unbiased cases.

Case (i) Biased ($p \neq q \neq \frac{1}{2}$)

$$q_z = \frac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^z}{\left(\frac{q}{p}\right)^a - 1} \qquad p_z = \frac{\left(\frac{q}{p}\right)^z - 1}{\left(\frac{q}{p}\right)^a - 1}$$

$$d_z = \frac{a}{q-p}\left[\frac{\left(\frac{q}{p}\right)^z - 1}{\left(\frac{q}{p}\right)^a - 1}\right] + \frac{z}{q-p}$$

Case (ii) Unbiased ($p = q = \frac{1}{2}$)

$$q_z = \frac{a-z}{a} \qquad p_z = \frac{z}{a}$$

$$d_z = z(a-z)$$

Note:- If the initial capital is doubled to $2z$ and $2(a-z)$ in unbiased case $d_{2z} = 2z\,2(a-z) = 4z(a-z) = 4d_z$

$$\boxed{d_{2z} = 4d_z}$$

Ex:- ~~If $p = \frac{1}{2}, q = \frac{1}{2}, z = 1, a = 500$ then~~

(1) If $p = \frac{1}{2}, q = \frac{1}{2}, z = 1, a = 500$ then find d_z

Sol: $d_z = z(a-z)$

$= 1(500-1) = 499$

(2) A gambler has 2 rupees, he bets 1 rupee at a time and wins 1/- with probability of $\frac{1}{2}$, He stops playing if he looses 2/- or wins

Ans: related markov chain.

i) What is the transition probability matrix of the related markov chain.

ii) What is the probability that he has loss money at the end of ~~5 plays~~ 5 plays

iii) What is the probability that game continuous more than ~~7 plays~~ 7-plays

Sol: Let X_n represents the amount of points left with the gambler

i.e. $X_n = \{0, 1, 2, 3, 4, 5, 6\}$

1) Transition probability matrix

$$
P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array}
\begin{array}{c}
\begin{array}{ccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\end{array}
$$

Probability before starts the game

$p^{(0)} = [0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0]$

Probability after 1st game

$p^{(1)} = p^{(0)}P = [0 \quad \frac{1}{2} \quad 0 \quad \frac{1}{2} \quad 0 \quad 0 \quad 0]$

$p^{(2)} = p^{(1)}P = [\frac{1}{4} \quad 0 \quad \frac{1}{2} \quad 0 \quad \frac{1}{4} \quad 0 \quad 0]$

$$p^{(3)} = p^{(2)}P = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & 0 & \frac{3}{8} & 0 & \frac{1}{8} & 0 \end{bmatrix}$$

$$p^{(4)} = p^{(3)}P = \begin{bmatrix} \frac{3}{8} & 0 & \frac{5}{16} & 0 & \frac{1}{4} & 0 & \frac{1}{16} \end{bmatrix}$$

$$p^{(5)} = p^{(4)}P = \begin{bmatrix} \frac{3}{8} & \frac{5}{32} & 0 & \frac{9}{32} & 0 & \frac{1}{8} & \frac{1}{16} \end{bmatrix}$$

$$p^{(6)} = p^{(5)}P = \begin{bmatrix} \frac{29}{64} & 0 & \frac{7}{32} & 0 & \frac{13}{64} & 0 & \frac{1}{8} \end{bmatrix}$$

$$p^{(7)} = p^{(6)}P = \begin{bmatrix} \frac{29}{64} & \frac{7}{64} & 0 & \frac{27}{128} & 0 & \frac{13}{128} & \frac{1}{8} \end{bmatrix}$$

ii) The probability that the player has loss the money at the end of 5 plays is $\frac{3}{8}$

iii) The probability that game Continues more than 7 plays

$$= \frac{7}{64} + \frac{27}{128} + \frac{13}{128} = \frac{27}{64}$$

4/10/16

UNIT-3.

Continous Probability Distributions.

[Normal, Exponential and Gama distrib

* Normal distribution:

The distributions binomial and poisson are discrete distributions where as the normal distribution is the continous distribution in which the variate can take all the possible values within the given range. It is also known as Guassion distribution. The normal distribution is the limiting from of binomial distribution. Here the values of 'n' are large and values of p and q are very small.

i.e, $n \to \infty$ and p (or) $q \to 0$

Defination: A random variable "X" is said to be a normal distribution. If its probability density function is given by $\boxed{f(x) = f(x, \sigma, \mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}$

where x lies b/w $-\infty$ to $+\infty$ and μ lies b/w $-\infty$ ...

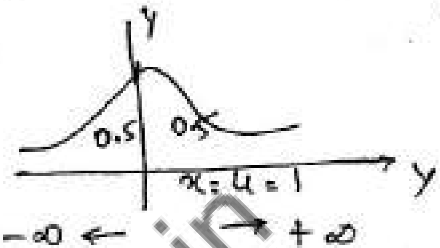
where $\sigma > 0$

Scanned with CamScanner

* Normal Curve :

The curve represting in normal distribution is known as Normal curve and the total area bounded by thats the curve by "x" axis i.e ⇒ 1. This curve has maximum value at $x = \mu$, and it extends both sides indefinitely and doesnot touches the horizontal line.

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



Constants of Normal distribution (ND) :

* To find mean of ND :

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

f(x) in N.d formula.

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

in eqⁿ (1

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

$\mu + \mu = b$

Scanned with CamScanner

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

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

let $z = \frac{x-b}{\sigma}$

$x - b = z\sigma$

$\boxed{x = z\sigma + b}$

$dx = \sigma \cdot dz + 0$

$\boxed{dx = \sigma dz}$

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

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

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

$$= \frac{1}{\sqrt{2\pi}} \left[\int_{0}^{\infty} z\sigma \, e^{-z^2/2} dz + b \int_{-\infty}^{\infty} e^{-z^2/2} d \right.$$

$$= \frac{1}{\sqrt{2\pi}} \left[0 + 2b \int_0^{\infty} e^{-z^2/2} \cdot dz \right]$$

$$= \frac{1}{\sqrt{2\pi}} (2b) \left[\int_0^{\infty} e^{-z^2/2} dz \right]$$

Scanned with CamScanner

$$\Rightarrow \frac{1}{\sqrt{2\pi}} 2b \times \sqrt{\pi/2}$$ Here we know that $\int_0^\infty e^{-z^2/2} dz = \sqrt{\frac{\pi}{2}}$

$$= \frac{1}{\cancel{\sqrt{2}}\cancel{\sqrt{\pi}}} \cancel{2}b \times \frac{\cancel{\sqrt{\pi}}}{\cancel{\sqrt{2}}} \Rightarrow b$$

$\boxed{\mu = b}$ $\boxed{b = \mu}$

* <u>Variance of the Normal distribution:</u>

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

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

Let $\mu = b$

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

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

let $z = \frac{x-b}{\sigma}$

$x - b = z\sigma$

$\boxed{dx = \sigma dz}$

$$\Rightarrow \frac{1}{\cancel{\sigma}\sqrt{2\pi}} \int_{-\infty}^{\infty} (z\sigma)^2 \cdot e^{-z^2/2} \times \cancel{\sigma} dz$$

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

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

Scanned with CamScanner

$$\frac{\sigma^2}{\sqrt{2\pi}} \times \cancel{2}\int_{-\infty}^{\infty} \underbrace{z^2 e^{-z^2/2}}_{\text{Even function}} dz$$

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

let $z^2 = 2t \Rightarrow z = \sqrt{2t}$

$\cancel{2}z\,dz = \cancel{2}\,dt$

$z\,dz = dt$

put $t=0 \Rightarrow t$

$z = 0 \Rightarrow t$

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

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

$$= \frac{\sqrt{2}\,\sigma^2}{\sqrt{\pi}}\int_0^{\pi} \sqrt{2t} \cdot e^{-2t/2}\,dz$$

$$= \frac{\sqrt{2}}{\sqrt{\pi}}\sigma^2\int_0^{\pi} \sqrt{2t} \cdot e^{-t} \cdot dt$$

$$= \frac{\sqrt{2}\times\sqrt{2}\,\sigma^2}{\sqrt{\pi}}\int_0^{\pi} \sqrt{t} \cdot e^{-t} \cdot dt$$

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

Scanned with CamScanner

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

Gama formula

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

$$= \int_0^{\infty} e^{-t} \cdot t^{3/2-1}\, dt$$

$$\cancel{\frac{2\sigma^2}{\sqrt{\pi}}} \times \cancel{\frac{\sqrt{\pi}}{2}}$$

$$\Gamma 3/2 - \Gamma 1/2 - 1$$

$$\cancel{\frac{\sigma^2}{\sqrt{\pi}}} \times \cancel{\sqrt{\pi}}$$

formula: $\boxed{\Gamma n+1 = n \Gamma n}$

$$\underline{\underline{\sigma^2}}.$$

$$\Gamma 1/2 + 1 = 1/2 \Gamma 1/2$$

formula: $\Gamma 1/2 = \sqrt{\pi}$]

Variance of Normal distribution $\boxed{=\sigma^2}$

$$x = dz\ \Gamma 3/2 = \tfrac{1}{2}\sqrt{\pi}$$

* Median of Normal distribution:

$$\int_a^b f(x)\, dx = \int_a^m f(x)\, dx = \int_m^b f(x)\, dx = 1/2$$

Let M is the median of the normal distribution

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

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

We know that N.D $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma}}$

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

Scanned with CamScanner

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

$\Rightarrow$ Consider $\int_{-\infty}^{\mu} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2} dx$

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

To change the limits:

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

$x - \mu = \sigma z$

$dx = \sigma dz$

$\frac{x-\mu}{\sigma} = z \rightarrow (1)$

let $x = \mu$ in (1)

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

$\Rightarrow z = 0$

~~$z = -\infty$~~

$z = -\infty$

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

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

By Symmetric property the $\frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-z^2/2} dz$

$$= \frac{1}{\sqrt{2\pi}} \times \sqrt{\pi/2}$$

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

$$= 1/2$$

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

Scanned with CamScanner

$$\Rightarrow \int_{-\infty}^{\mu} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \int_{\mu}^{M} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2} dx =$$

$$+ \int_{\mu}^{N} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2} dx = \frac{1}{2}.$$

$$= \int_{\mu}^{n} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2} dx = \cancel{\frac{1}{2}} - \cancel{\frac{1}{2}} \Rightarrow 0$$

W.K.T $\int_a^b f(x)\, dx = 0 \Rightarrow a = b$.

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

$$\Rightarrow \boxed{M = \mu} //$$

Median of the normal distribution $\boxed{M = \mu}$ //

Mode of the Normal distribution :

Mode is the value of x for which the $f(x)$ is mean is mode is solution of $f'(x) = 0 \Rightarrow f''(x) < 0$

$$f(x) > 0 \quad , \quad f(x) = 0 \quad , \quad f'(x) < 0$$

By defination, we have :

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

Scanned with CamScanner

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

$$= \frac{1}{\sigma\sqrt{2\pi}} \, e^{-1/2\left[\frac{x-\mu}{\sigma}\right]^2} \times d/dx \; -1/2\left[\frac{x-\mu}{\sigma}\right]^2$$

$$= 1/\sigma\sqrt{2\pi} = e^{-1/2\left[\frac{x-\mu}{\sigma}\right]^2} \times -\cancel{1/2} \times \cancel{2}\left[\frac{x-\mu}{\sigma}\right]$$

$$= -1/-\sigma\sqrt{2\pi} \; e^{-1/2\left[\frac{x-\mu}{\sigma}\right]^2} \times -\left[\frac{x-\mu}{\sigma}\right]$$

$$= f(x) \times -\left[\frac{x-\mu}{\sigma}\right]$$

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

let $f'(x) = 0 \Rightarrow -\left[\frac{x-\mu}{\sigma}\right] \; f(x) = 0$

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

$$x - \mu = 0 \quad x = \mu$$

$$f''(x) = -d/dx \left[\left(\frac{x-\mu}{\sigma}\right)\right] f(x)$$

$$f'(x) = -d/dx \left[\left(\frac{x-\mu}{\sigma}\right) f(x)\right]$$

u v

$$-\left[\left(\frac{x-\mu}{\sigma}\right) d/dx \; fx + f(x) \; d/dx \left[\frac{x-\mu}{\sigma}\right]\right]$$

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

Let $x = \mu$ in (2)

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

$$\boxed{f''(x) = \tfrac{1}{\sigma} f(x) < 0}$$

Hence, ~~Mode~~ $x = \mu$ is mode of Normal distributi

$$\therefore \boxed{\text{Mean} = \text{Median} = \text{Mode} = \mu.}$$

* Mean ~~derivative~~ deviation of mean of the Normal distribution

By defination mean deviation about mean:

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

W.K.T N.d formula is $= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$.

$$\Rightarrow \int_{-\infty}^{\infty} |x - \mu| \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2} dx.$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} |x - \mu|\, e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2} dx.$$

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

$$x - \mu = \sigma z$$

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

Scanned with CamScanner

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} |\sigma z| \, e^{-1/2 \, z^2} \times \sigma \, dz$$

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

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

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

$$\therefore z^2 = 2t$$

$$\cancel{2} \cdot z \, dz = \cancel{2} \, dt \quad = z \, dz = dt$$

$$\Rightarrow \sqrt{2}/\sqrt{\pi} \cdot \sigma \int_{0}^{\infty} e^{-t} \cdot dt$$

$$= \sqrt{2}/\sqrt{\pi} \, \sigma \times \left(-e^{-t}\right)_{0}^{\pi}$$

$$= \sqrt{2}/\sqrt{\pi} \, \sigma - \left[e^{-\pi} - e^{0}\right]$$

$$\sqrt{2}/\sqrt{\pi} \, \sigma - (0 - 1)$$

$$= \sqrt{2/\pi} \, \sigma \times 1 = \sqrt{2}/\sqrt{\pi} \, \sigma$$

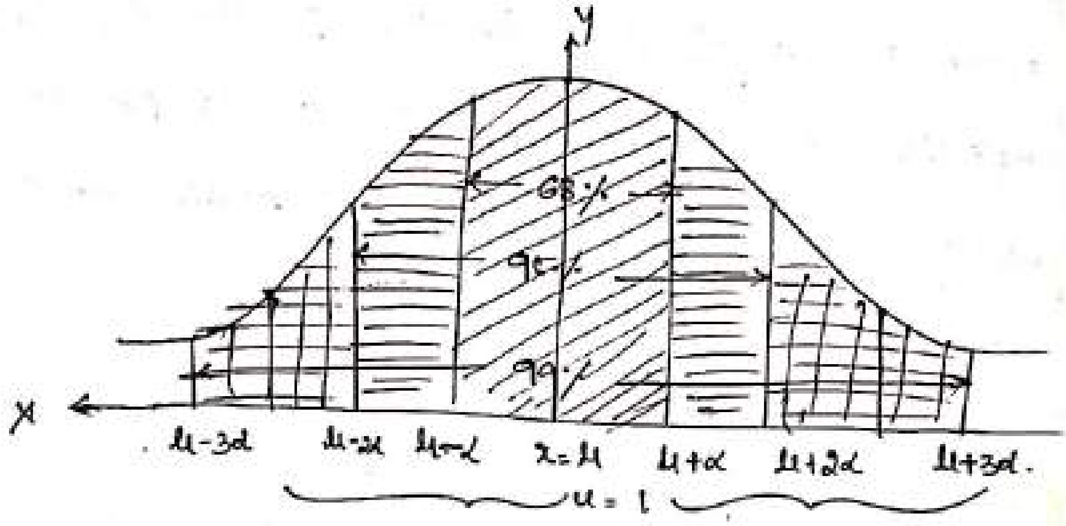
$$= \boxed{\frac{4}{5} \sigma} = \frac{4\sigma}{5} \text{ (appr}$$

Hence of mean deviation of mean of n.d is $4/5$ times of standard deviation (appro)

Scanned with CamScanner

17/9/19.

Characteristics of Normal distribution:



The mean, median, mode for the normal distribution are identical

* The curve is bell shaped and symmetrical about the line x i.e. $x=\mu$

* The area under the curve above x-axis. is from $-\infty$ to $+\infty$

* The area of the Normal curve b/w $(\mu-\sigma)$ & $(\mu+\sigma)$ is 68% i.e. $P[\mu-\sigma < x < \mu+\sigma] = 68\%$.

* The area of the normal curve b/w $(\mu-2\sigma)$ & $(\mu+2\sigma)$ is 95% i.e $P[\mu-2\sigma < x < \mu+2\sigma] = 95\%$.

* The area of Normal curve b/w $(\mu-3\sigma)$ and $(\mu+3\sigma)$ is 99% i.e, $P[\mu-3\sigma < x < \mu+3\sigma] = 99\%$.

Scanned with CamScanner

Formulas : [Normal curve]

* The probability that the normal curve variate X w mean μ and standard deviation σ lies b/w two specific values x_1 and x_2 with $x_1 \le x_2$ can be obt using area under the standard normal curve as follows:

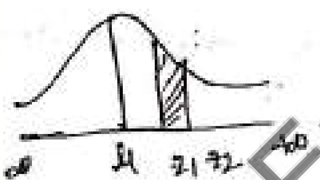
Step 1 :

Case 1: Both z_1 and z_2 are positive (or) Negative

(1) $z_1 > 0$, $z_2 > 0$

$\downarrow$

$$P(x_1 \le X \le x_2) = P(z_1 \le z \le z_2)$$



$$P(x_1 < X \le x_2) = \boxed{|A(z_2) - A(z_1)|}$$

$\boxed{z_1 > 0, z_2 > 0}$

+ve

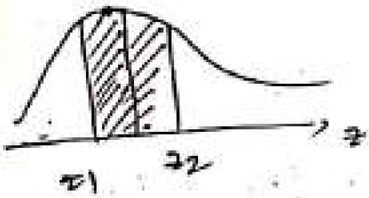
$\rightarrow$ $z_1 < 0$, $z_2 < 0$

$$P(x_2 \le X < x_1) = \boxed{|A(-z_2) - A(z_1)|}$$

$\boxed{z_1 < 0, z_2 < 0}$

Case 2: If $z_1 < 0$ and $z_2 > 0$

(2) $P[x_1 \le X \le x_2] = |A(z_2) + A(z_1)|$



$< \; >$

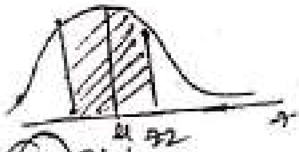
$+$

$> \; > \; -$

$< \; <$

$\ge \; \le$

Case 3: If $z_1 > 0$ and $z_2 < 0$



$P[x_2 \le x \le x_1] = |A(z_2) - A(z_1)|$

(3)

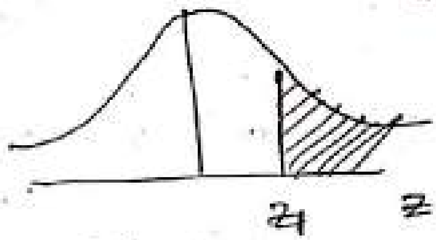
Step (2):

To find $P(z > z_1)$:

Case (1): If ~~$z_1 < 0$~~ $z_1 > 0$ then

$P(z > z_1) = 0.5 - A(z_1)$

(1) [Since $P(z<0) = P(z>0) = 1/2$]



Case (2): If $z_1 < 0$ then

$P[z < z_1] = 0.5 + A(z_1)$

(2)

Case (3): To find $P[z < z_1] = 1 - P[z > z_1]$

(3)

Scanned with CamScanner

Step ③ :

Case 1 :- If $z_1 > 0$ then $P(z < z_1) = 1 - P(z > z_1)$

$= 1 - [0.5 - A(z_1)]$

Case 2 :- If $z_1 \le 0$ then $P(z < z_1) = 1 - P(z > z_1)$

$= 1 - [0.5 + A(z_1)]$

18/9/19.

σ^2

①

1) For a normally distributed variate with mean (μ) 1 and standard deviation (σ) 3. find the probabilities that:

i) $\overset{x_1}{3.43} \le x \le \overset{x_2}{6.19}$.

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

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

Sol: Given that :

Mean $(\mu) = 1$.

Standard deviation $(\sigma) = 3$.

i) $\overset{x_1}{3.43} \le x \le \overset{x_2}{6.19}$.

$P[x_1 \le x \le x_2] = P[z_1 \le z \le z_2]$

$x_1 = 3.43$ and $x_2 = 6.19$

Scanned with CamScanner

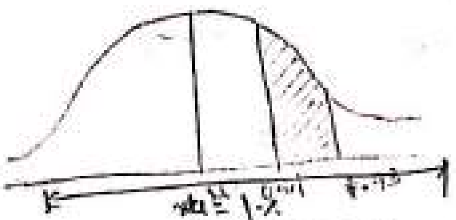
$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{3.43 - 1}{3} = 0.81 \Rightarrow z_1 > 0$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{6.19 - 1}{3} = 1.73 \Rightarrow z_2 > 0$$

We know that:

$z_1 > 0$ and $z_2 > 0$.

$$P[x_1 \le X \le x_2] = P[z_1 \le Z \le z_2]$$



[From Table]

$$= |A(z_2) - A(z_1)|$$

$$\Rightarrow |A(1.73) - A(0.81)|$$

$$= |0.4582 - 0.2910|$$

$$= 0.1672$$

ii) $-1.43 < x \le 6.19$.

$$P[x_1 \le X \le x_2] = P[z_1 \le Z \le z_2]$$

$x_1 = -1.43$, $x_2 = 6.19$. $z_1 \le 0$.

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{-1.43 - 1}{3} = -0.81 \Rightarrow z_1 < 0$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{6.19 - 1}{3} = 1.73 \Rightarrow z_1 > 0$$

we know that:-

$z_1 < 0$ and $z_2 > 0$.

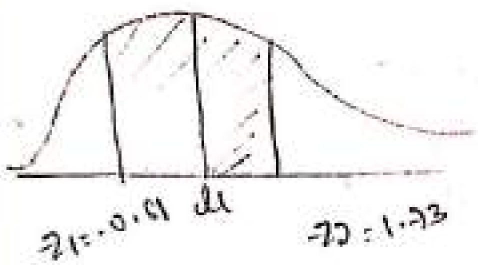
$$P[x_1 \le X \le x_2] = |A(z_2) + A(z_1)|$$

formula

$[A(-z) = -A(z)]$

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

$$= |.4582| + |(-0.\overset{2910}{\cancel{2881}})|$$



$$= 0.4582 + 0.2910$$

$$= \underline{0.7492}.$$

② x is the normal variate with mean 30 and the standard deviation 5. find i) $P[26 \le x \le 40]$

ii) $P[x \ge 45]$

Sol: Mean $(\mu) = 30$

$(\sigma) = 5$

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

$$= P(x_1 < x < x_2) = P(z_1 < z < z_2)$$

$$x_1 = 26, \quad x_2 = 40$$

Scanned with CamScanner

$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{26-30}{5} = -0.8 \Rightarrow z_1 < 0$

$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{40-30}{5} = 2 \Rightarrow z_2 > 0.$

Now, we know that :

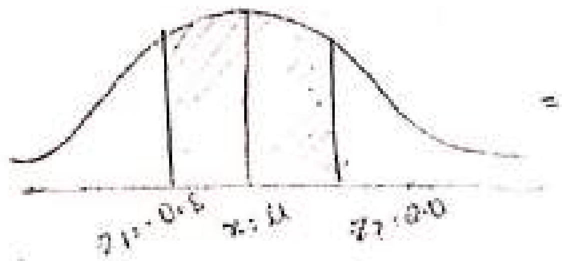
$z_1 < x < z_2.$

$z_1 < 0 \; \& \; z_2 > 0.$

$z_1 < 0$ and $z_2 \geq 0$

$P[x_1 \leq x \leq x_2] = |A(z_2) + A(z_1)|$

$= |A(2) + A(-0.8)|$



$= |0.4772| + |\cancel{-0.2881}|$ 0.2881

$= 0.4772 + \cancel{0.2881}$ 0.2881

$= \cancel{0.7653}$ 0.7653

(ii) $P(x \geq 45)$.

$\frac{x-\mu}{\sigma}$ $\frac{45-30}{5}$

$P(z \geq z_1) = P(x \geq 45)$

$z \geq 3.$

If $z > 0 \Rightarrow P(z > z_1) = 0.5 - A(z_1)$ $3.03.$

$= 0.5 - A(\cancel{3.03})$ 3.0

Here $z_1 = \frac{x_1 - \mu}{\sigma}$

$= \frac{45-30}{5}$

$\boxed{z_1 = 3}$

$= 0.5 - 0.4987$

$= 0.0013$

Scanned with CamScanner

③ The mean and standard deviation of marks obtained by 1000 students in an examination are respectively 34.5 and 16.5. Assuming the normality of distribution find the approximate no. of students expected obtaining marks b/w 30 and 60.

Sol: Given that:

$$\text{Mean}(\mu) = 34.5$$

$$\text{Standard deviation}(\sigma) = 16.5$$

Given $x_1 = 30$ and $x_2 = 60$.

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{30 - 34.5}{16.5} = -0.27 < 0 \Rightarrow z_1 < 0$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{60 - 34.5}{16.5} = 1.54 > 0 \Rightarrow z_2 > 0.$$

We know that:

$$z_1 < 0 \text{ and } z_2 > 0$$

$$P[x_1 < x < x_2] = |A(z_2) + A(z_1)|$$

$$P[30 \leq x \leq 60] = |A(1.54) + A(-0.27)|$$

$$= |0.4382| + |0.1064|$$

$$= 0.4382 + 0.1064$$

0.5446

Scanned with CamScanner

∴ The 1000 numbers of students who get marks b/w 30 and 60 are $\Rightarrow$ 0.5446×1000

$$= \underline{544.6}$$

∴ Hence, ~~544~~ 545 students get marks b/w 30 and 60.

(1) Imp. M.Q.2 In a normal distribution 7% of items are under 35% and 89% are under 63. determine the mean and variance of the distribution.

Sol: Given that:

Let μ is the mean and σ is the standard deviation of the normal curve.

* $P(x \leq 35) = 7\% \Rightarrow 7/100 = 0.07$

* $P(x \leq 63) = 89\% \Rightarrow 89/100 = 0.89$

* $P(x > 63) = 1 - P(x \leq 63) = 1 - 0.89$

$$= 0.11$$

Scanned with CamScanner

$\therefore x_1 = 35 \Rightarrow \frac{x_1 - \mu}{\sigma} = \frac{35 - \mu}{\sigma} = -z_1$ (say) $\rightarrow$ (1)

$x_2 = 63 \Rightarrow \frac{x_2 - \mu}{\sigma} = \frac{63 - \mu}{\sigma} = z_2$ (say) $\rightarrow$ (2)

$\Rightarrow$ from figure:

$P[0<z<z_1] = \cancel{0.39}\ \cancel{0.43} = z_1$

$0.5 - A(z_1)$

$= 0.43$

$\boxed{z_1 = 1.48}$ from Table where 0.43 value given

$P[0<z<z_2] = 0.5 - A(z_2)$

$= 0.39$

$\boxed{z_2 = 1.23}$

We have from eq (1) $\frac{35 - \mu}{\sigma} = -1.48 \rightarrow$ (3)

from eq (2) $\frac{63 - \mu}{\sigma} = 1.23 \rightarrow$ (4)

$\therefore 35 - \mu = -1.48\sigma$

$(-)\ 63 \ (+)\ \mu = (-)\ 1.23\sigma$

$\cancel{-}28 = \cancel{-}2.71\sigma$

$\sigma = 28/2.71 \Rightarrow 10.332$

Scanned with CamScanner

Now $\sigma^2 = 10.332^2 = 106.75$ // $\boxed{\sigma^2 = 106.75}$ //

Put $\sigma = 10.332$ in eqⁿ (3)

$$\frac{35-\mu}{\sigma} = \frac{-1.48}{1} \rightarrow (3)$$

$$35-\mu = -1.48\sigma$$

$$35-\mu = -1.48 \times (10.332)$$

$$\cancel{35}\,\mu = \cancel{+}\,1.48 \times 10.332 + 35$$

$$\boxed{\mu = 50.3}$$ //

$$\sigma^2 = 106.75$$

$$\sigma = \sqrt{106.75}$$

$$\boxed{\sigma = 10.332}$$ //

Mean $(\mu) = 50.3$ //

Variance $(\sigma^2) = 106.75$ //

Standard deviation $(\sigma) = 10.332$ //

(4) In a normal distribution 35% of items under 45 and 8% are over ~~64%~~ 64. find the mean & variance of distribution.

Sol:

Given:

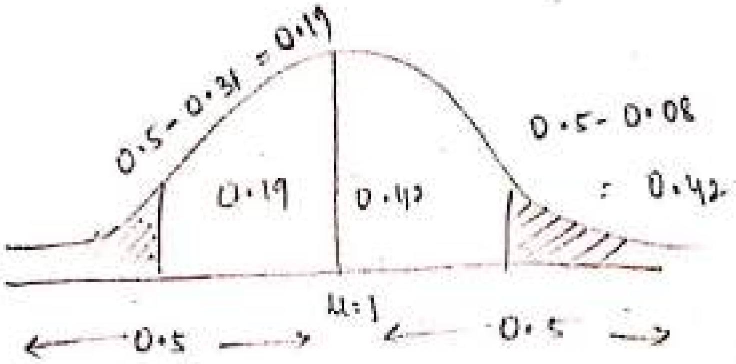
35% of → 45 under

8% of → 64 over

Scanned with CamScanner

* $P(x \le 45) = 31\% \Rightarrow 31/100 \Rightarrow 0.31$

* $P(x \ge 64) = 8\% \Rightarrow 8/100 \Rightarrow 0.08$



$x_1 = 45 \Rightarrow \frac{x_1 - \mu}{\sigma} = \frac{45 - \mu}{\sigma} = -z_1$ (say)

$x_2 = 64 \Rightarrow \frac{x_2 - \mu}{\sigma} = \frac{64 - \mu}{\sigma} = z_2$ (say)

from figure :

$P[0 < z < z_1] = 0.5 - A(z_1)$

$= 0.19$

$z_1 = 0.50$

$P[0 < z < z_2] = 0.5 - A(z_2)$

$= 0.42$

$z_2 = 1.41$

Scanned with CamScanner

$\sigma^2 = 98.94$

$\sigma = \sqrt{98.94}$

$\boxed{\sigma = 9.94}$

Mean (μ) = 50

Variance (σ^2) = 98.94

Standard deviation (σ) = 9.94

01/2/19.

1) X is a normal variate, find the area A

① To left side of $z = -1.78$

② To the right side of $z = -1.45$

③ Corresponding to $-0.8 \le z \le 1.53$

④ To the left of $z = -2.52$ & right of $z = 1.83$

1) Given that,

To left of $z = -1.78 < 0$ $\quad 0.5 - A(z_1)$

$P(z < z_1)$ if $z_1 < 0$

Scanned with CamScanner

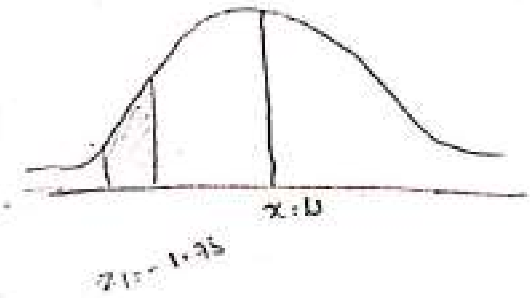
Scanned with CamScanner

formula Req. area = $= 0.5 - A(z_1)$

$= 0.5 - A(1.78)$ (from Tables)

$= 0.5 - 0.4625$

$\Rightarrow 0.0375$



ii) $P(Z > z_1)$, $z_1 = -1.45 < 0$

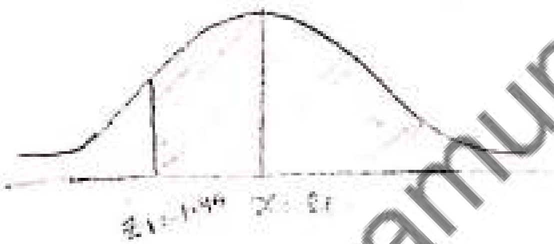
$P(Z > z_1)$, $z_1 < 0$

formula required area $\Rightarrow 0.5 + A(z_1)$

$= 0.5 + A(1.45)$

$= 0.5 + 0.4265$

$\Rightarrow 0.9265$



iii) Corresponding $P[-0.8 \le Z \le 1.53]$

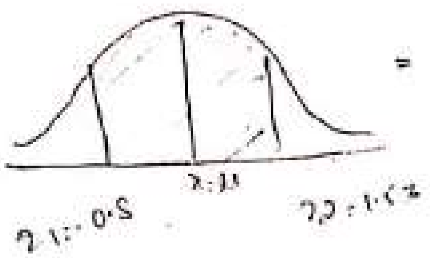
$P[-z_1 \le z \le z_2]$

$z_1 = -0.8 < 0$, $z_2 = 1.532$

$z_1 < 0$ & $z_2 > 0$ $\Rightarrow$ Required Area $= |A(z_1) + A(z_2)|$

$= |A(-0.8) + A(1.53)|$

$= A(0.8) + A(1.53)$



$= 0.2881 + 0.4370$

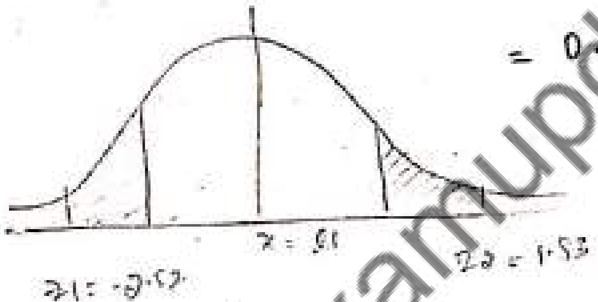
$= \underline{0.7251}$

iv) To the left $z = -2.52$ and right of $z = 1.83$

$P(z < z_1),\ z_1 < 0 + P(z > z_2),\ z_2 > 0$

Required area $= 0.5 - A(z_1) + 0.5 - A(z_2)$

$= 0.5 - A(2.52) + 0.5 - A(1.83)$



$= 0.5 - 0.4941 + 0.5 - 0.4664$

$= \underline{\underline{0.0395}}$

* In a sample of 1000 cases the mean of a certain test is 14. and standard deviation is 2.5. Assuming the distribution to be normalised, find

i) how many students score between 12 and 1

ii) " " " " above 18

iii) " " " " below 18

Sol:

i) $P[x_1 \le x \le x_2] = P(12 \le x \le 15)$

$x_1 = 12$ and $x = 15$

Given that Mean $(\mu) = 14$

s.d $(\sigma) = 2.5$

$z_1 = \frac{x - \mu}{\sigma} = \frac{12 - 14}{2.5} = -2/2.5 \Rightarrow -0.8 < 0$

$z_1 < 0$

$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{15 - 14}{2.5} = 1/25 \Rightarrow 0.4 > 0$

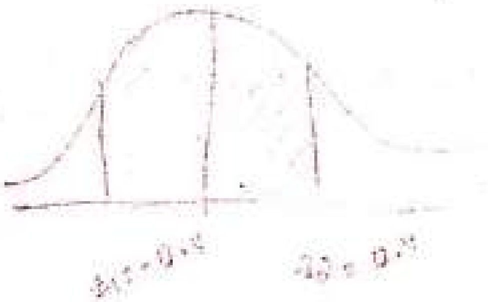
$z_2 > 0$

$z_1 < 0$ and $z_2 > 0$ then;

$P[x_1 \le x \le x_2] = P[z_1 \le z \le z_2] = |A(z_1) + A(z_2)|$

$P[12 \le x \le 15] = P[-0.8 \le z \le 0.4]$

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



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

$= 0.2881 + 0.1554$

$= 0.4435$

Scanned with CamScanner

ii) Above 18 :

$x_1 = 18$

Mean $(\mu) = 14$ s.d $(\sigma) = 2.5$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{18-14}{2.5} = 1.6$$

$z_1 > 0$

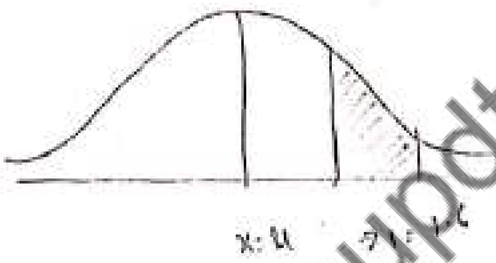
$P(z > z_1)$, $z_1 > 0$

Required area (A) $= 0.5 - A(z_1)$

$= 0.5 - A(1.6)$

$= 0.5 - 0.4452$

$= 0.0548.$



iii) Below 18 :

$x_1 = 18$, $\mu = 14$, $\sigma = 2.5$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{18-14}{2.5} = 4/2.5 = 1.6$$

$-1.6 < 0$ [Given that condition is below]

$P(x < x_1) = P(z < z_1) + z_1 > 0$

Required area $= 0.5 + A(z_1)$

$= 0.5 + A(1.6)$

$= 0.5 + 0.448$

$\Rightarrow 0.945$

Scanned with CamScanner

Imp

- The marks obtained in maths by 1000 students is normally distributed with mean 78% and standard deviation is 11%. determines

$\frac{2r}{60} = 0\ z_1$

i) How many students got above 90%

ii) What was the highest mark obtained by the lowest 10% of students ?

iii) Within what limits what did the 90% of the students like.

Sol: Given data:

$\mu = 78\% \Rightarrow 78/100 = 0.78$

$\sigma = 11\% \Rightarrow 11/100 = 0.11$

① $P[x > x_1] = P[x > 90\%]$

$= P[x > 90/100 = 0.9]$

$P[x > 0.9]$

$x_1 = 0.9$

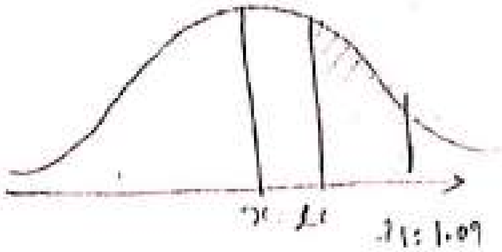
$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{0.9 - 0.78}{0.11} \Rightarrow 1.09 > 0$

$z_1 > 0$

Scanned with CamScanner

$P(z > z_1)$ and $z_1 > 0$

Required Area $= 0.5 - A(z_1)$



$= 0.5 - A(1.09)$

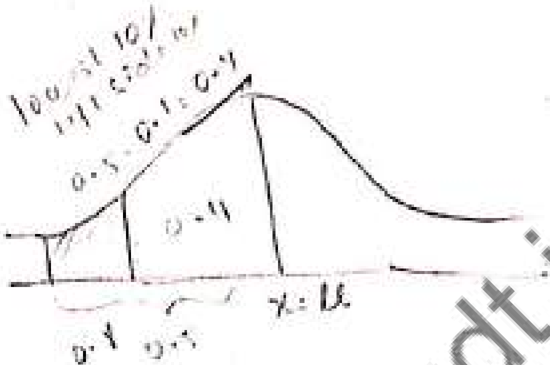
$= 0.5 - 0.362$

$= 0.1379$

$\therefore$ No. of students are $0.1379 \times 1000 \Rightarrow 137.9$

$= 138$ stu

iii)



$\underline{0.4} = 0.5 - $ Area from 0 to z_1

$z_1 = -1.29$

$-1.29 = \frac{x_1 - 0.78}{0.01}$

$-1.29 \times 0.01 = x_1 - 0.78$

$z_1 = \frac{x_1 - \mu}{\sigma}$

$-1.29 \Rightarrow \frac{x_1 - 0.78}{\sigma}$

$-1.29 \times 0.11 = x_1 - 0.78$

$-0.14 = x_1 - 0.78$

$x_1 = -0.78 + 0.14$

$x_1 = 0.6392$

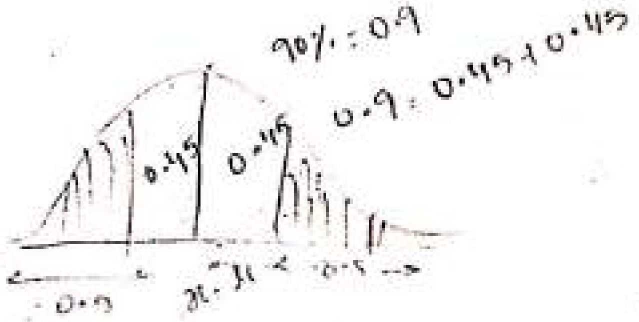
$= 0.6392 \times 100$

$= 63.92 \simeq 64\%$

Scanned with CamScanner

64% of students got highest Marks.

iii) Middle of 90% of region:



$0.4 < 0.5$

Middle of 90% of region corresponding to 0.9 area

i.e = $0.45 + 0.45$

→ Leaving 0.05 on both sides then area from 0 to

z_1 = 0.5 - Area of 0 to z_1

$0.45 = 0.5$ - area of $(0 - z_1)$

$z_1 = -1.65$ (from the normal distribution Table)

$0.45 = 0.5$ - area of $(0 - z_2)$

$z_2 = 1.65$

$z_1 < 0$ and $z_2 > 0$

Scanned with CamScanner

$$Z_1 = \frac{x-\mu}{\sigma} \Rightarrow -1.65 = \frac{x_1 - 0.78}{0.11}$$

$$x_1 = 0.5985$$

$$x_1 = \underline{59.85\%}$$

$$Z_2 = \frac{x-\mu}{\sigma} \Rightarrow 1.65 = \frac{x_2 - 0.78}{0.11}$$

$$x_2 = 0.9615$$

$$x_2 = 96.15\%$$

* ~~If~~ the masses of 300 students are normally distributed with mean 68 kgs and standard deviations 3 kgs. How many students have masses ~~i)~~ greater than 72 kgs

ii) b/w 65 and 71 kgs
x_1 x_2

Sol: Given:

Mean (μ) = 68 kgs

s.d (σ) = 3 kgs.

(i) Greater than 72 kgs.

$x_1 > 0$.

$$Z_1 = \frac{x_1 - \mu}{\sigma} = \frac{72 - 68}{3} = 4/3 = 1.33$$

Scanned with CamScanner

* $P(z>0) \Rightarrow P(1.33>0) \Rightarrow 0.5 - A(z_1)$

$= 0.5 - 1.33$

$= 0.5 - 0.4082$

$= \underline{0.0918}$

ii) B/w $\underline{65}$ and $\underline{71}$ kgs :-

$$P(x_1 \leq x \leq x_2) = P(65 \leq x \leq 71)$$

$$z_1 = \frac{65-68}{3} \Rightarrow -1 < 0 \quad (z_1 < 0)$$

$$z_2 = \cancel{46}\ \frac{71-68}{3} \Rightarrow 1 > 0 \quad (z_2 > 0)$$

$$P(z_1 < z \leq z_2| = |A(z_1) + A(z_2)|$$

$$= |A(-1) + A(1)|$$

$$= A(1) + A(1)$$

$$= 0.3413 + 0.3413$$

$$= \underline{0.6826}$$

Scanned with CamScanner

* Suppose the weight of 800 male students are normally distributed with mean $\mu = 140$ pounds and s.d is (σ) 10 pounds. Find the no. of students whose weight are:

1) b/w $\overset{x_1}{138}$ and $\overset{x_2}{148}$ pounds

2) More than 152 pounds

Sol: Given data:

Mean $(\mu) = 140$ pounds

Standard deviation $(\sigma) = 10$ pounds.

1) b/w 138 and 148 pounds:

$$p(x_1 \le x \le x_2) = p(138 \le x \le 148)$$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{138 - 140}{10} = -0.2 < 0 \quad z_1 < 0$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{148 - 140}{10} = 0.8 > 0 \quad z_2 > 0$$

$$p(z_1 \le z \le z_2) = |A(z_1) + A(z_2)|$$

$$= |A(-0.2) + A(0.8)|$$

$$= A|0.2| + A|0.8|$$

$$= 0.0793 + 0.2881$$

$$= \underline{\underline{0.3674}}$$

Scanned with CamScanner

2) More than 152 pounds :

$z_1 = \frac{x_1 - \mu}{\sigma}$ $z_1 = \frac{152 - 140}{10}$

$z_1 = 12/10 = 1.2.$

$P(z > 0) = P(1.2 > 0) \Rightarrow 0.5 - A(z_1)$

$= 0.5 - A(1.2)$

$= 0.5 - 0.3849$

$= \underline{0.1151}$

* The mean inside diameter of a sample of 200 washers produced by a machine is 0.500 with sd 0.005 cm. The purpose of which these washers are intensioned is a maximum tolerance in the diameter 0.495 to 0.505 cm. Otherwise the washers are considerd defective. determine the % of defective washers produced by the machine Assumed the diameters are normally distributed.

Sol: We are given that

μ = Mean of the inside diametres

$= 0.500$ cm

Scanned with CamScanner

$\sigma = 0.005$ cm.

→ Two tolerance limits of non defective tubes are was

$x_1 = 0.495$, $x_2 = 0.505$ cm.

then P(defective washers) $= 1 - p$[non defective]

$= 1 - p(x_1 \le x \le x_2)$

$= 1 - p[0.495 < x < 0.5$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{0.495 - 0.500}{0.005} = -1 < 0 .$$

$z_1 < 0$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{0.505 - 0.500}{0.005} = +1 > 0$$

$z_2 > 0$

formulas

$z_1 < 0$ & $z_2 > 0 \Rightarrow 1 - p(z_1 \le z \le z_2)$

$= 1 - p[-1 < z < +1]$

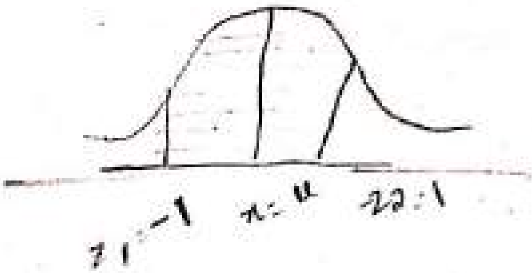
$= 1 - |A(z_1) + A(z_2)|$

$= 1 - A(-1) + A(1)|$

$= 1 - [0.3413 + 0.3413$

$\Rightarrow 0.3174$ //

Scanned with CamScanner



Percentage of defective washing = 31.74%.

* In an examination it is laid down that a student passes of his ~~case~~ secures 40% or more. He is placed in the 1st, 2nd and 3rd division accordingly as he secures 60% or more marks; b/w 50 and 60% are second class and b/w 40 & 50% respectively. He gets the distinction incase he secures 75% or more. It is noticed from the results that 10% of students failed in the examination where 5% of them obtained the distinction. Calculate the % of students placed in 2nd division.

Sol: let μ = Mean and σ = standard deviation

If percentage of students of failed in exam = 10%.

i.e $P(x<40) = 10\% = 0.1$

Scanned with CamScanner

Homework Problems

1) Given that the mean height of students in a class i.e, 158 cms with standard deviation of 20 cms. find how many students heights lies b/w 150 and 170 cms. if there are 100 students in the class.

Sol: Given that:

$$\text{Mean } (\mu) = 158 \text{ cms}$$

$$\text{s.d } (\sigma) = 20 \text{ cms}$$

→ B/w 150 cm and 170 cm.

$$P(x_1 \le x \le x_2) = P(z_1 \le z \le z_2)$$

$$x_1 = 150 \text{ and } x_2 = 170$$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{150 - 158}{20} = -0.4 \qquad z_1 < 0$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{170 - 158}{20} = 0.6 \qquad z_2 > 0$$

We know that: $z_1 < 0$ and $z_2 > 0$

Then $P(z_1 \le z \le z_2) = |A(z_2) + A(z_1)|$

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

$= 0.2258 + 0.1554$

$= 0.3812$

Scanned with CamScanner

There are 100 students in class = $0 \cdot 3812 \times 100$

$= 38 \cdot 12 = 38$

$\therefore$ 38 members of students lies b/w 150 & 170 cms

(2) The mean height of students in a college is 155 cms and standard deviation is 15. What is the propability that the mean height of 36 students is less than 157 cms.

Sol: Given data:

$$\text{Mean}(\mu) = 155 \text{ cms}$$

$$\text{S.d}(\sigma) = 15 \text{ cms}$$

Scanned with CamScanner

③ 1000 students have written an examination. The mean of test is 35 and standard deviation is 5. Assuming the distribution is normal, find.

① How many students marks lie b/w 25 and 40?

② How many students got more than 40?

③ How many students got below 20?

④ How many students got more than 50?

Sol: Given data:

Mean (μ) = 35

Standard deviation (σ) = 5.

i) B/w 25 and 40:

$$P(x_1 \le x \le x_2) = P(z_1 \le z \le z_2)$$

$x_1 = 25$ and $x_2 = 40$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{25-35}{5} \Rightarrow -2 < 0 \quad \boxed{z_1 < 0}$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{40-35}{5} \Rightarrow 1 > 0 \quad \boxed{z_2 > 0}$$

Scanned with CamScanner

W.K.T;

$z_1 < 0$ and $z_2 > 0$

then $[P[z_1 \le z \le z_2] = |A(z_2) + A(z_1)|$

$= |A(z_2)| + A(z_1)|$

$= A(-0.2) + A(1)$

$= A|0.2| + A|1|$

$= 0.0793 + 0.3413$

$= \underline{0.4206}$

$\Rightarrow 0.4206 \times 1000 = 420.6$

No. of students got marks b/w 25 & 40 are $\underline{421}$.

(2) More than 40 &

$(x > z_1)$

$$z_1 = \frac{x_1 - \mu_1}{\sigma} = \frac{40-35}{5} = 5/5 = 1$$

$\underline{z_1 > 0}$.

$P(z > z_1), z_1 > 0$

Required area $= 0.5 - A(z_1)$

$= 0.5 - A(1)$

$= 0.5 - 0.3413$

$= 0.1587$

Scanned with CamScanner

$\Rightarrow 0.1587 \times 1000 \Rightarrow \underline{158.7}$

$\therefore$ No. of studes who got more than 40 are = $\underline{159}$

③ Below 20 Marks :-

$x = 20, \mu = 35, \sigma = 5.$

$z_1 = \frac{x-\mu}{\sigma} = \frac{20-35}{5} = -3, \quad \underline{z_1 < 0}.$

$P(z > z_1)$ and $z_1 < 0$.

Required area $= 0.5 + A(z_1)$

$= 0.5 + A(-3)$

④ More than 50 Marks :-

$x = 50, \mu = 35, \sigma = 5.$

$z_1 = \frac{50-35}{5} = 15/5 = 3, \quad z_1 > 0.$

$P(z > z_1)$ and $z_1 > 0$ The area $= 0.5 - A(z_1)$

$= 0.5 - A(3)$

$= 0.5 - 0.4987$

$\Rightarrow 0.0013$

Scanned with CamScanner

$\Rightarrow 0.0013 \times 1000 \Rightarrow 1.3 \approx 1.$

$\therefore$ 1 student got more than 50 marks.

(4) Problem:

In a test on 200 electrical bulbs. It was found that the life of a particular mean was normally distributed with an average life of 2040 hours and s.d of 40 hours. Estimate the no. of bulbs likely to burn for i) More than 2140 hours

ii) b/w 1920 and 2080 hours

iii) less than 1960 hours

Sol:- Given that :

$$\text{Mean } (\mu) = 2040$$

$$\text{s.d } (\sigma) = 40 \text{ hrs}$$

i) More than 2140 hours :

$$[x > z_1]$$

$$z_1 = \frac{x-\mu}{\sigma} = \frac{2140 - 2040}{40} = \frac{100}{40}$$

$$= 2.5 > 0$$

Scanned with CamScanner

$p(z > z_1)$, $z > 0$

Required Area $= 0.5 - A(z_1)$

$\Rightarrow 0.5 - A(2.5)$

$= 0.5 - 0.4938$

$= \underline{0.0062}$

ii) b/w 1920 and 2080 hours :-

$p(x_1 \le x \le x_2) = p(z_1 \le z \le z_2)$

$x_1 = 1920$ and $x_2 = 2080$

$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{1920-2040}{40} = -3$

$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{2080-2040}{40} = 1$

WKT $z_1 < 0$ and $z_2 > 0$

then $P[z_1 \le z \le z_2] = |A(z_2) + A(z_1)|$

$= |A(1) + A(3)|$

$= 0.3413 + 0.4987$

$= \underline{0.84}$

Scanned with CamScanner

iii) less than 1960 hours :

$x = 1960$ $\quad z_1 = \frac{x-\mu}{\sigma} = \frac{1960-2040}{40} \Rightarrow$

$p(z_1 \quad z) = \quad z_1 > 0$

* If x has the binomial distribution with mean 25 and probability of success 1/5. find $P[x < \mu - 2\sigma]$ where μ and σ^2 are mean and variance of the distribution.

Sol: Given :

$$\text{Mean}(\overset{np}{\mu}) = 25$$

$$\text{Variance }(npq) = ?$$

$$P = 1/5$$

$$\therefore np = 25 \left[\cancel{1/5}\right] = 25/5$$

$$n(1/5) = 25$$

$$n = 25 \times 5$$

$$\boxed{n = 125}$$

$$q = 1 - p$$

$$= 1 - 1/5$$

$$\boxed{q = 4/5}$$

Scanned with CamScanner

Variance (σ^2) = npq

$= 125 \times 1/5 \times 4/5$

$= 20$

$\sigma^2 = 20$

$\sigma = \sqrt{20}$

$\boxed{\sigma = 4.47}$

$\rightarrow \mu - 2\sigma$

$= 25 - 2 \times (4.47)$

$\boxed{\mu - 2\sigma = 16.06}$

$\Rightarrow P[x < \mu - 2\sigma]$

$= P[x < 16.06]$

$= P\left[z < \frac{x-\mu}{\sigma}\right]$

$= P\left[z < \frac{16.06 - 25}{4.47}\right]$

$= P[z < -2]$

Here $P[z < z_1]$ if $z_1 < 0$

formula $0.5 - A(z_1)$

$= 0.5 - 0.4772$

$= 0.0228$

Scanned with CamScanner

we know that $z = \frac{x-\mu}{\sigma}$

Let z_1 and z_2 be the values of z, corresponding to x_1 and x_2 of x respectively.

$$P[x_1 < x < x_2] = P[z_1 < z < z_2] = \int_{z_1}^{z_2} \phi(z)\, dz$$

Case 2:

When $p \neq q$

for large n, we can approximate the binomial curve by the normal curve and calculated the probability.

→ for any sucess x, real intrevel is $x - 1/2, x + 1/2$.

Hence z_1 must be corresponding to the lower limit of x_1 and z_2 be the upper limit of x_2.

$$z_1 = \frac{(x_1 - 1/2) - \mu}{\sigma} = \frac{(x_1 - 1/2) - np}{\sqrt{npq}}$$

$$z_2 = \frac{(x_2 + 1/2) - \mu}{\sigma} = \frac{(x + 1/2) - np}{\sqrt{npq}}$$

Scanned with CamScanner

* **Normal Approximation to the Binomial distribution**

The normal distribution can be used to approximate the binomial distribution [B.d]. Suppose the number of success x ranges from x_1 to x_2, then probability of getting x_1 to x_2 success is given by:

$$\sigma = \sum_{r=x_1}^{x_2} {}^nC_r p^r q^{n-r}$$

for large n, the calculation of binomial property is very difficult. In such causes the binomial curve cannot be replaced by the Normal curve. In this process, we consider two cases:

Case 1: If $p = q = 1/2$. when n is not large, binomial distribution (bd) can be approximated by Normal distribution [Nd].

$\therefore$ Mean of binomial distribution $= np$

Standard deviation $(\sigma) = \sqrt{npq}$

Hence, for the corresponding normal distribution μ and σ are more.

Scanned with CamScanner

① Find the probability that the guess work. A student will correctly answer 25 to 30 questions in a multiple choice quiz consisting of 80 questions. Assume that in each question with 4 choices, only 1 choice is correct and the student has no knowledge of the subject.

Sol: Here $p = 1/4$, $n = 80$

$$q = 1 - 1/4$$

$$q = 3/4.$$

$$\text{Mean}(\mu) = np = n \times \left(\frac{1}{4}\right)$$

$$= 80 \times 1/4$$

$$\mu = 20$$

$$\sigma = \sqrt{npq}$$

$$= \sqrt{80 \times (1/4) \times (3/4)}$$

$$\sigma = \sqrt{15}$$

$$\boxed{\sigma = 3.87}$$

$$\cancel{z_1 = \frac{x_1 - \mu}{\sigma}} \Rightarrow \cancel{\frac{25 - 20}{3.87}} = \cancel{1.29}$$

$\cancel{z_1 = 0}$

Scanned with CamScanner

Scanned with CamScanner

$\therefore\ p[25 \le x \le 30]$

$= x_1 \le x \le x_2.$

Here $x_1 = 25$ and $x_2 = 30$.

$$\rightarrow z_1 = \frac{(x_1 - 1/2) - \mu}{\sigma} = \frac{(x_1 - 1/2) - np}{\sqrt{npq}}$$

$$= \frac{(25 - 1/2) - 20}{3.87} = \underline{1.16}$$

$$\rightarrow z_2 = \frac{(x_2 + 1/2) - \mu}{\sigma} = \frac{(x_2 + 1/2) - np}{\sqrt{npq}}$$

$$= \frac{[30 + 1/2] - 20}{3.87} = \underline{2.71}$$

$\Rightarrow\ p(z_1 \le z \le z_2),\ z_1 > 0$ and $z_2 > 0$

Required Area $= |A(z_2) - A(z_1)|$

$= A(2.71) - A(1.16)$

$= 0.4966 - 0.3770$

$= \underline{0.1196}$

Scanned with CamScanner

③ Find the probability of getting an even number face 3 to 5 times in throwing 10 dice together.

Sol: p = probability of getting an even number on face of dice $= \frac{1}{6} + \frac{1}{6} + \frac{1}{6}$

$$= \frac{3}{6}$$

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

$$q = 1 - p$$

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

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

$$n = 10$$

$$\therefore \mu = np \Rightarrow 10\left(\frac{1}{2}\right) = \boxed{\mu = 5}$$

$$\sigma = \sqrt{npq} = \sqrt{10 \times \left(\frac{1}{2}\right) \times \left(\frac{1}{2}\right)} = 1.58$$

$\sqrt{npq}$.

Formula: $\int_{z_1}^{z_2} \phi(z)\, dz$

$P(x_1 \le x \le x_2)$

$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{3 - 5}{1.58} \Rightarrow \cancel{-1.26}$ $P(3 \le x \le 5)$

$x_1 = 3$ & $x_2 = 5$

$\cancel{z_2 = \frac{x_2 - \mu}{\sigma} = \frac{5-5}{1.58}}$

Scanned with CamScanner

$$z_1 = \frac{[x_1 - 1/2] - \mu}{\sigma} = \frac{[3-0.5] - 5}{1.58} \Rightarrow -1.58$$

$z_1 < 0$

$$z_2 = \frac{(x_2 - 1/2] - \mu}{\sigma} = \frac{(5+0.5] - 5}{1.58} \Rightarrow 0.316$$

$z_2 > 0.$

$$\Rightarrow \int_{z_1}^{z_2} \phi(z)\, dz = \int_{-1.58}^{0.32} \phi(z)\, dz$$

$$= P[z_1 \leq z \leq z_2]$$

$z_1 < 0$ and $z_2 > 0$.

$$\Rightarrow |A(z_1) + A(z_2)|$$

$$= A(\cancel{\ }-1.58| + A|0.32|$$

$$= 0.\cancel{\ }4429 + 0.1256$$

$$= \cancel{0.5685}$$

$$\underline{\underline{0.5685}}$$

③ 8 coins are tossed together find the probability of getting 1 to 4 in a single toss

Sol: Given that:

P = probability of getting 1 to 4

$\boxed{P = 1/2}$ $q = 1-p = 1 - 1/2$ $\boxed{q = 1/2}$

Scanned with CamScanner

$n = 8$

$\mu = np \Rightarrow 8[1/2] \Rightarrow 4$

$\sigma = \sqrt{npq} = \sqrt{4 \times (1/2)} \Rightarrow \underline{1.41}$

$\rightarrow \int_{z_1}^{z_2} \phi(z) \cdot dz$

$z_1 = \frac{(x_1 - 1/2) - \mu}{\sigma} = \frac{(1 - 1/2) - 4}{1.41} = \underline{-2.48}$

$z_2 = \frac{(x + 1/2) - \mu}{\sigma} = \frac{(4 + 1/2) - 4}{1.41} = \underline{0.35}$

$\int_{z_1}^{z_2} \phi\, z \cdot dz$

$= \int_{-2.48}^{0.35} [P(z_1) < z < z_2)$

$z_1 < 0$ and $z_2 > 0 = |A(z_1) + A(z_2)|$

$= A(-2.48) + A(0.35)|$

$= 0.4934 + 0.1368$

$= 0.6302$

Scanned with CamScanner

UNIT-5

①

Testing of Hypothesis.

i. Test of Hypothesis [for large samples]:

① One Tail Test

② Two tail test

③ Type-I and Type-II error.

④ Hypothesis conferring one and two metho

ii. Small Samples:

① t-distribution, f-distribution, χ^2 distribut

② students t-test

③ " " f-test.

Hypothesis:

To estimate the value of the parameter from a statistic of a sampling distribution. We need to decide whether to accept (or) reject a statement about the parameter. This statement is called as hypothesis and the decision making procedure about hypothesis is called Testing of hypothesis.

Scanned with CamScanner

②

This is one of the most useful aspects of statistical procedure.

Procedure for Testing of a Hypothesis :-

Step 1 : There is two types of hypothesis

i) Null hypothesis (H_0) :

A null hypothesis is a hypothesis which asserts that there is no significance difference between the statistic and population parameter.

If any difference is observed that may be due to the fluctuations in sampling. We usually defined it has a definite statement about the population parameter.

→ It is denoted by H_0.

→ It is the form $\boxed{H_0 : \mu = \mu_0}$ i.e., when even we test whether the procedure is better than the other, we assume that there is no difference b/w them.

Scanned with CamScanner

Alternate Hypothesis (H_A):

Any hypothesis which contradicts the null hypothesis is called an alternate hypothesis. It is denoted by H_1.

It is of the form $\boxed{H_1 : \mu \neq \mu_0}$ either $\{\mu < \mu_0 \text{ (or) } \mu > \mu_0\}$

↓

Two tailed test

→ One Tailed test → $\begin{cases} H_1 : \mu < \mu_0 & \text{[left tailed 1·H]} \\ H_1 : \mu > \mu_0 & \text{[right tailed 1·H]} \end{cases}$

Level of Significance:

It is confidence which we accept or reject the null hypothesis i.e, it is the max. probability which we are willing to risk an error in rejecting H_0, when it is true.

→ It is denoted by α.

→ It is also known as size of the Test.

Eg: for 1% → level of significance there is only one case in 100%, that H_0 is rejected when it is true

Scanned with CamScanner

Step 4: Test statistic:

There are several test of significance for both small samples and large samples.

→ We select test samples depending on the nature of information given.

Step 5: Result (or) Conclusion:

We compute the values of appropriate statistic which fails in the acceptance (or) rejection region depends on the value of α.

→ If computed value (or) calculated value is less than critical (or) tabular values, we reject H_0. otherwise we accept H_0.

Imp

* Errors:

The difference b/w calculated value and actual value (tabular value).

Scanned with CamScanner

We have two types of Error :

i) Type-I Error.

ii) Type-II Error.

i) Type-I Error :

Reject null hypothesis when it is true, when null hypothesis is true, but the difference is significant & hypothesis is rejected, then type-2 error is made, the probability of making type-1 error is denoted by α.

→ It is also called α - error.

→ The probability of making correct decision is $(1-\alpha)$.

ii) Type-II Error :

Accept null hypothesis, when it is wrong, if null hypothesis falls, but it is accepted by test then the error is called type-II error or β error.

Scanned with CamScanner

6

Critical Region:

A region corresponding to the sample statistic "t" in sample space "s" which leads to the rejection of H_0 is called Critical Region (or) Rejection region.

Acceptance region:

The region which leads to the

Critical values (or) Significant values:

The value of the test statistic which separates the critical region and acceptance region is called Critical value.

→ The values depends on the level of the significance and Alternate Hypothesis.

Scanned with CamScanner

One Tail Test and Two Tailed Test :

(7)

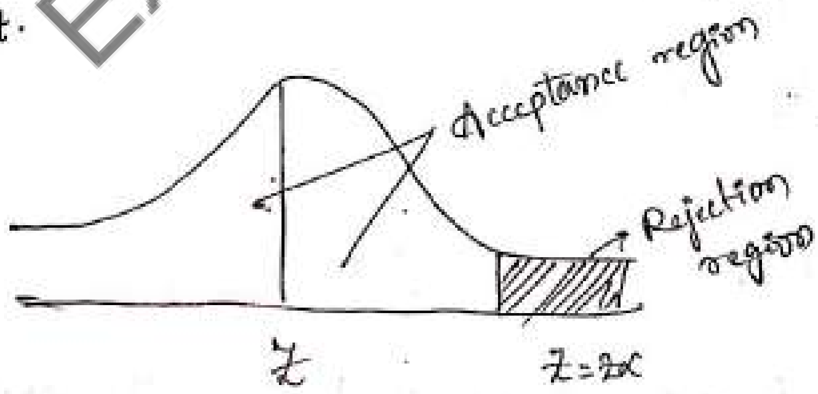
(i) One Tailed Test :

If the alternate hypothesis (H_1) in a test of statistical hypothesis be said to be one tailed. Then test is one tailed test (either left or right).

For Example: To test whether the population mean $\mu = \mu_0$. It is also known as single Tailed Test.

Right tailed Alternate hypothesis :

If $H_1 : \mu > \mu_0$, then A.H. said to be right tailed Test.

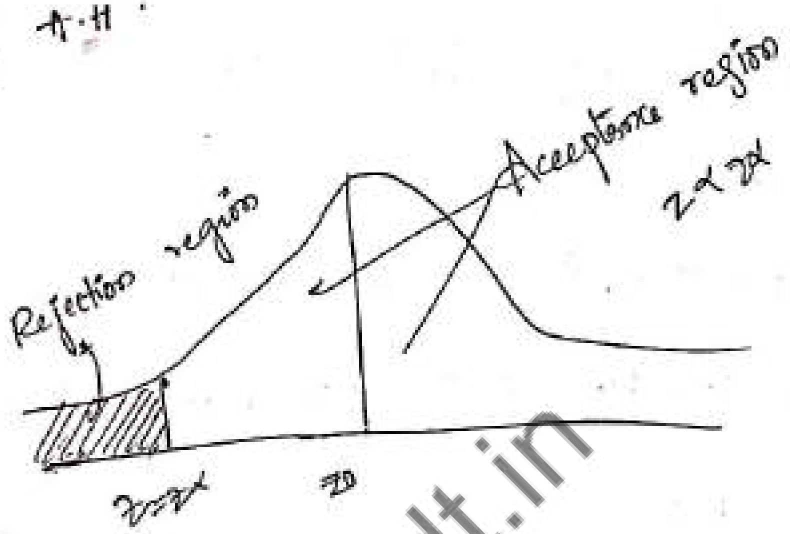


In a right tailed test, the critical region, $Z > Z_\alpha$ lies entire the in the right tailed of the sample distribution of sample mean $\bar{x}$

with area equal to the level of significance. (8)

Left tailed T.H:

If $H_1: \mu < \mu_0$, then T.H said to be left tailed T.H.



In a left tailed test $Z < Z\alpha$ lies entirely in the left tail of the sampling distribution of sample mean $\bar{x}$ with area equal to the level of significance.

→ Two Tailed Test:

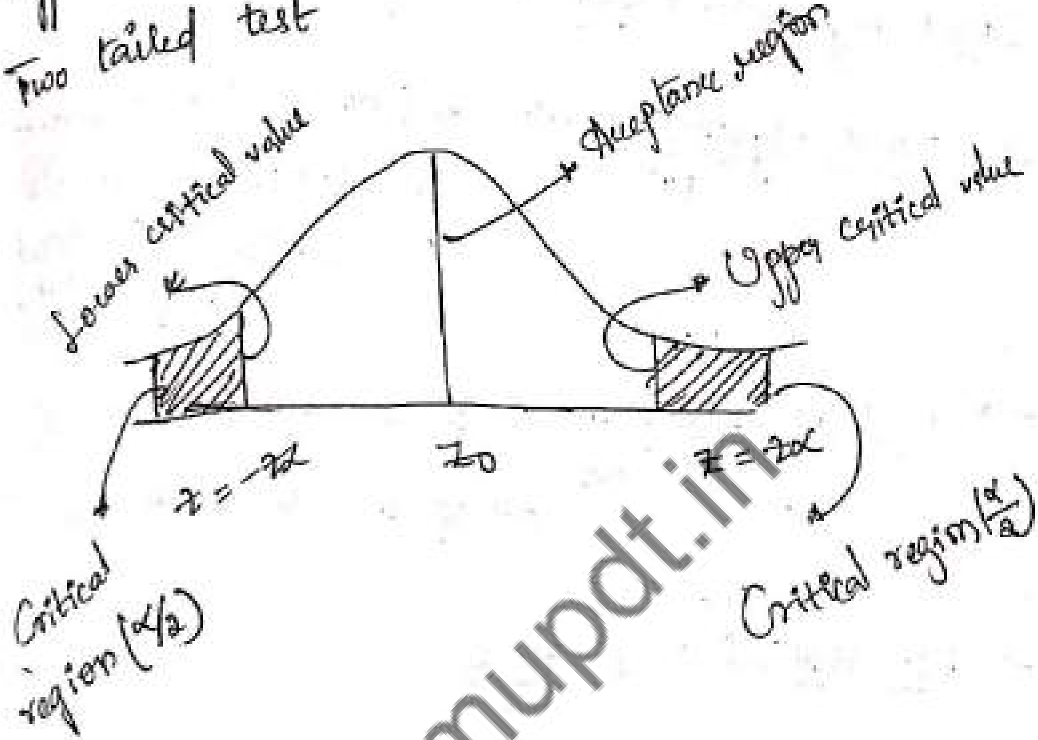
Suppose we want to test the null hypothesis (H_0) against the alternative hypothesis is eve

Scanned with CamScanner

that $\mu_0 \neq \mu_1$, that is

$$H_1 = \mu_0 \neq \mu_1 \text{ i.e; } \mu > \mu_0 \text{ (or) } H_1; \mu < \mu_0.$$

→ If alternate hypothesis is in a test of statistical hypothesis is two tail, then the test is called Two tailed test



→ Critical Area under right tail = Critical area under left tail $= \frac{1}{2}$ [Total area] $= \frac{\alpha}{2}$

∴ With critical statistic

Scanned with CamScanner

9/11

Test of Significance for Large Samples.

* Procedure for test of significance for large sampl or procedure for testing of hypothesis:

→ Null hypothesis (H_0): Define (or) setup $(H_0) = \mu = \mu_0$

→ Alternate Hypothesis (H_1): $H_1 = \mu \neq \mu_0$ → Two tailed t.

(or) $\mu > \mu_0$ → Right Tailed
$\mu < \mu_0$ } → One Tailed t.t.
↳ left tailed.

→ Level of Significance (α): It is usually we choose 5% of level of significance.

→ Test statistic: $z = \frac{\bar{x} - \mu}{\sigma}$

→ Conclusion (or) Result: $Z_{cal} < Z_{Tab}$, $|z| < z$.

Then Null hypothesis is Accepted otherwise test is Rejected.

Large Sample:

If the sample size $N \geq 30$ then we called it is called Large sample.

Scanned with CamScanner

Table [Critical values of z]

Level of significance (α).	1% (or) 99%	5% (or) 95%	10% (or 9
Critical value for two tailed test	$\lvert z_\alpha \rvert = 2.58$	$\lvert z_\alpha \rvert = 1.96$	$\lvert z_\alpha \rvert = 1$
Critical value for Right tailed test.	$\lvert z_\alpha \rvert = 2.33$	$\lvert z_\alpha \rvert = 1.645$	$\lvert z_\alpha \rvert = 1$
Critical values for left tailed test.	$\lvert z_\alpha \rvert = -2.33$	$\lvert z_\alpha \rvert = -1.645$	$\lvert z_\alpha \rvert = -1.2$

Problems:

1) A coin was tossed 960 times and returned heads 183 times. Test the hypothesis that the coin is unbaised. Use a 0.05 level of significance.

Sol: Here $n = 960$. , Probability of getting head $= 1/2$ (p)

$q = 1/2$.

$\mu = np$

$= 960 \times (1/2)$

$\boxed{\mu = 480}$

Scanned with CamScanner

* Standard deviation $(\sigma) = \sqrt{npq}$

$$= \sqrt{960 \times \frac{1}{2} \times \frac{1}{2}}$$

$$\boxed{\sigma = 15.49}$$

No. of success $(x) = 183$.

i) Null hypothesis: $H_0 : \mu = \mu_0$

The coin is unbaized.

ii) Alternate Hypothesis:

$(H_1) : \mu \neq \mu_0$

The coin is baized.

iii) Level of significance:

$\alpha = 0.05 = 5\%$

iv) Test statistic: $Z = \frac{x - \mu}{\sigma} = \frac{183 - 480}{15.49} = -19.17$

$|Z cal| = |-19.17| = 19.17$

v) Result:

If $\alpha = 0.05$ from table, Then $Z_\alpha = 1.96$

$|Z| cal = 19.17$ $\quad Z table(\alpha) = 1.96$

So $Z cal > Z table$. $\Rightarrow 19.17 > 1.96$

So, Test is Rejected @ 5% level of significance

$\therefore$ The coin is baized.

Scanned with CamScanner

3) A die is tossed 960 times and it falls with 5 upwards 184 times. It is the die unbaized at a level of significance of 0.01.

Sol: Given that:

$n = 960$

Probability of getting die $(P) = 1/6$

$q = 1 - p = 1 - 1/6$

$\boxed{q = 5/6}$

$\mu = np$

$= 960 \times 1/6$

$\boxed{\mu = 160}$

$\sigma = \sqrt{npq}$

$= \sqrt{960 \times 1/6 \times 5/6}$

$\boxed{\sigma = 11.54}$

$x = 184$.

i) Null hypothesis: $H_0 : \mu = \mu_0$

The ~~coin~~ die is unbaized.

ii) Alternate hypothesis: $(H_1) : \mu \neq \mu_0$

The die is baized.

iii) Level of significance:

$\alpha = 0.01 = 1\%$

Scanned with CamScanner

iv) Test statistic :

$$Z = \frac{x - \mu}{\sigma} = \frac{184 - 160}{11.54} = 2.07$$

$Z_{cal} = 2.07$, $Z_{table} = 2.58$.

v) Result :

If $Z_\alpha = 2.58$ from table

$Z_{cal} = 2.07$.

$Z_{cal} < Z_{tab}(\alpha)$.

$2.07 < 2.58$

$\therefore$ so, the test is accepted @ 1% of level of significance.

$\therefore$ The die is unbaized.

Model-2

Testing of hypothesis for single Mean of Large samples :

Let a random sample of size $n \geq 30$ has a sample mean $\bar{x}$ and μ be the population mean. Also the population mean (μ) has a specifi value "μ_0".

Scanned with CamScanner

Null hypothesis: $(H_0): \bar{x} = \mu$

i.e, There is no significant difference b/w sample mean and Population mean.

Alternate hypothesis (H_1):

$$(H_1): \bar{x} \neq \mu.$$

(or) $H_1: \bar{x} > \mu$

$H_1: \bar{x} < \mu.$

ii) Level of significance:

i) Test Statistic:

$$\bar{x} = \sigma/\sqrt{n} \qquad z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Result:

$|z| < z_\alpha.$

where z_α = critical value of z (or) Tabular value of z

Note:

The confidence limits of given single mean $\bar{x} \pm z_\alpha \left(\frac{\sigma}{\sqrt{n}}\right)$.

where $\frac{\sigma}{\sqrt{n}}$ is standard Error.

Scanned with CamScanner

→ A sample of 400 items is taken from a population whose s.d is 10. The mean of sample is 40. Test whether the sample has come from a population with mean 38. Also calculate 95% confidence interval

Sol: Given that:

$n = 400 \geq 30$ [Large sample]

$\sigma = 10$

$\mu = 38$

$\bar{x} = 40$.

i) Null Hypothesis: $H_0 : \bar{x} = \mu$.

ii) Alternate hypothesis:

$H_1 : \bar{x} \neq \mu$

There is significant difference b/w $\bar{x}$ and μ.

iii) Level of significance:

$\alpha = 5\%$.

iv) Test statistics:

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{40 - 38}{\frac{10}{\sqrt{400}}} = \underline{\underline{4}}$$

$|Z|_{cal} = 4$

② → According to the normal established for a mechan[illegible] aptitude test persons who are 18 years old. have an average height of 73.2 with a s.d of (σ) 8.6. If [illegible] randomly selected persons of that is overaged $\bar{x}$ 76.7 Test the hypothesis $\boxed{\mu = 73.2}$ against the alternate hypothesis $\mu > 73.2$. @ level of significance 0.01

Sol: Given:

$n = 4$

$\mu = 73.2 \quad \bar{x} = 76.7$

$\sigma = 8.6$

i) Null Hypothesis: $H_0 : \bar{x} = \mu$.

There is no difference b/w sample mean & population mean

ii) Alternate Hypothesis: $H_1 : \mu > 73.2$.

[Right Tailed Test]

iii) level of significance:

$\alpha = 0.01 = 1\%$

iv) Test statistics: $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{76.7 - 73.2}{\frac{8.6}{\sqrt{4}}} = 0.81$

$|z|_{cal} = 0.81$

v)

Scanned with CamScanner

$Z_\alpha = 1.96.$

Conclusion:

$Z_\alpha = 1.96$, $Z_{cal} = 4$ @ 5%.

$|Z| = 4$

$Z_{cal} > Z_\alpha$.

$4 > 1.96$

Rejected Null Hypothesis (H_0)

∴ Therefore sample has not come from the population

To find confidence limits: @ $\alpha = 95\%$.

$$\Rightarrow \bar{x} \pm Z_\alpha \left[\frac{\sigma}{\sqrt{n}}\right]$$

$$\Rightarrow 40 \pm 1.96 \left[\frac{10}{400}\right]$$

$$\left[40 + 1.96 \left(\frac{10}{400}\right) \quad 40 - 1.96 \left[\frac{10}{400}\right]\right]$$

$$= [40.98, 39.02]$$

Scanned with CamScanner

i) Result: $Z\alpha = 2.33$ @ $\alpha = 1\%$.

$|Z|_{cal} = 0.81 < Z\alpha = 2.33$.

$Z < Z\alpha$. $\Rightarrow 0.81 < 2.33$.

Accept the Null Hypothesis.

$\therefore$ Sample had difference b/w sample & population mean

3) An ambulance service claims that it takes an average less than 10 minutes the reach its destination in emergency halts. A sample of 36 calls has a mean of 11 minutes and the variance of 16 minutes. Test the ~~client~~ claim at 0.05 level of significance

Sol: Given:

$n = 36$, $\bar{x} = 11$

$\sigma^2 = 16$

$\sigma = \sqrt{4}$

$\boxed{\sigma = 2}$

$\sigma^2 = 16$

$\sigma = \sqrt{16}$

$\sigma = 4$

$\mu = 10$

i) Null Hypothesis (H0): $\bar{x} = \mu$.

There is no significance difference b/w sample & population mean

Scanned with CamScanner

ii) Alternate Hypothesis :

$H_1 : \mu < 10$.

That is left tailed test

iii) level of significance :

$\alpha = 0.05 = 5\%$.

iv) Test statitis :

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{11-10}{\frac{4}{\sqrt{36}}} = \frac{11-10}{\frac{4}{\sqrt{6}}} = 1.5$$

$|Z_{cal}| = 1.5$

v) Result :

At $\alpha = 5\%$. $Z_{tab} = -1.645$

$Z_{cal} < Z_{\alpha} = 1.5 < -1.645$

The test is rejected.

Model No : 3

Test of significance for difference of Means

(or) Test for Equallity of two means :

let $\bar{x}, \bar{x_1}, \bar{x_2}$ be the sample means of two independent large random sample sizes - n_1, n_2

Scanned with CamScanner

drawn from the two populations having means of μ_1 and μ_2 and s.d's are σ_1 and σ_2. To test whether 2 populations means are equal.

1) Null hypothesis: $H_0 : \bar{x_1} = \bar{x_2}$

2) Alternate H: $H_1 : \bar{x_1} \neq \bar{x_2}$

3) Level of significance: α is given value

4) Test statistics:

$$Z = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \longrightarrow (1)$$

formula

If $\boxed{\sigma_1 = \sigma_2 = \sigma}$ then

$$Z = \frac{\bar{x_1} - \bar{x_2}}{\sigma\sqrt{1/n_1 + 1/n_2}} \longrightarrow 2^{nd} \text{ formula.}$$

5) Result:

$|Z|_{cal} < Z_\alpha \rightarrow$ Accept Test

$|Z|_{cal} > Z_\alpha \rightarrow$ Reject Test

Note: Confidence limits of difference of two Means:

$$(\bar{x} - \bar{y}) \pm Z_\alpha \cdot \text{standard error of } (\bar{x} - \bar{y})$$

Scanned with CamScanner

⑩ The means of two large samples of sizes of 1000 and 2000. Numbers are 67.5 and 68. inches respt. Can the samples be regarded as drawn from same population of s.d 2.5 inches.

Sol: Given:

let μ_1 and μ_2 be the means of two population

$$n_1 = 1000 \geq 30 \text{ (large)}$$

$$n_2 = 2000 \geq 30 \text{ (large)}$$

$\bar{x}_1 = 67.5$ inches $\quad \bar{x}_2 = 68$ inches.

$$\sigma = 2.5 \text{ inches.}$$

Then the formula is $\sigma_1 = \sigma_2 = \sigma = 2.5$

1) Null hypothesis:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{1/n_1 + 1/n_2}}$$

$$Z = \frac{67.5 - 68}{2.5\sqrt{1/1000 + 1/2000}}$$

$$Z_{cal} = 5.16$$

$$|Z|_{cal} = |5.16| = 5.16$$

Scanned with CamScanner

i) Null Hypothesis (H0) : $\bar{x_1} = \bar{x_2}$

The samples regarded as drawn from the same population

ii) A. Hypothesis : (H1) : $\bar{x_1} \neq \bar{x_2}$

The samples regarded are not drawn from the same population

iii) Level of S : $\alpha = 5\%$ (95%)

iv) Test. S : $Z = \underline{5.16}$

Given formula below.

v) Result :
$Z_\alpha = 1.96$, $Z_{cal} = 5.16$

$Z_{cal} > Z_\alpha$.

Rejected Null hypothesis at 5%.

→ Sample are not drawn for same population of S.d 2.5 inches.

Scanned with CamScanner

② The average marks scored by 32 boys ~~a~~ is 72 with a s.d (σ) of 8. while that for 36 girls is 70 with s.d of 6. Thus, this indicate that the boys performed better than girls that $\alpha = 0.05$.

Sol: Given:

let μ_1 and μ_2 be means of two populations.

i) Null. H: $H_0 : \mu_1 = \mu_2$

ii) A.H: $H_1 : \mu_1 \neq \mu_2$ $(\mu_1 > \mu_2)$

iii) level of S: $\alpha = 5\%$.

iv) T.S: $Z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$.

Here $n_1 = 32$, $n_2 = 36$

$\bar{x}_1 = 72$. $\bar{x}_2 = 70$, $\sigma_1 = 8, \sigma_2$

$$\Rightarrow \frac{72 - 70}{\sqrt{8^2/32 + 6^2/36}}$$

$$\boxed{Z = 1.15}$$

Scanned with CamScanner

Result:- $Z_{cal} = 1.15$

$Z_\alpha = 1.96$

$Z_\alpha > Z_{cal} \Rightarrow 1.96 > 1.15$

Accepted Null Hypothesis.

A researcher wants to know the intelligence of students in a girl. He selected 2 groups of students. in the first group, their 150 students having mean avg of 75 with sd of 50. In second group, their 250 students having mean avgque of 70 with sd of 20. Is there a significance difference b/w the means of two groups.

Sol:- Given:-

$n_1 = 150, \ n_2 = 250$

$\bar{x}_1 = 75, \ \bar{x}_2 = 70$

$\sigma_1 = 15, \ \sigma_2 = 20$

1) Null H:- $H_0: \mu_1 = \mu_2$.

The groups have been come from same population.

Scanned with CamScanner

ii) A.H: $H_1: \mu_1 \neq \mu_2$

The groups have not been came from same Popu

It is called Two tailed (because not mentioned any)

iii) α.O.S:

$\alpha = 5\%$

iv) T.S:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma^2/n_1 + \sigma^2/n_2}}$$

$$= \frac{75 - 70}{\sqrt{15^2/150 + 20^2/250}}$$

$|Z|cal = 2.83$

v) Conclusion:

Z_α from table $= 1.96$

$|Z|cal = 2.83$

$|Z|cal > Z_\alpha \Rightarrow 2.83 > 1.96$

Reject Null Hypothesis

The group have not drawn from same

Scanned with CamScanner

The reported nicotent in milligrams of 2 samples of tobacco where found to be as follows. find the standard error & confidence limits for difference b/w means of $\alpha = 0.05$.

Sample-A	24	27	26	23	25	-
Sample B	29	30	30	31	24	36

Sol: $n_1 = 5$ and $n_2 = 6$

$\bar{x} = \frac{\Sigma x_i}{N} = \frac{125}{5} = 25$

$\bar{y} = 180/6 = 30$

x_i	y_i	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
24	29	$(24-25)^2 = 1$	1
27	30	4	0
26	30	1	0
23	31	4	1
25	24	0	36
-	36	-	36
$\Sigma x_i = 125$	$\Sigma y_i = 180$	10	$\Sigma(y_i - \bar{y})^2 = 74$

Scanned with CamScanner

$$Z = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$ (σ is unknown)

s_1^2 and s_2^2 are variance of samples.

$\longrightarrow s_1^2 = \sum_{i=1}^{n} = \frac{(\bar{x_i} - \bar{x_1})^2}{n_1} \Rightarrow s_1^2 = \frac{10}{n_1} = 10/5 = 2$

$\longrightarrow s_2^2 = \; = \frac{(y_i - \bar{y_1})^2}{n_2} \Rightarrow s_2^2 \Rightarrow 74/n_2 = 74/6 = 12.$

$Z = 2.5$

standard error $= \sqrt{s_1^2/n_1 + s_2^2/n_2}$

$= \sqrt{2/5 + 12.3/6}$

$= \underline{1.56}$

Confidence limits :-

on 5% = 1.96 (for Z table)

$(\bar{x} - \bar{y}) \pm Z\alpha$ (SE of $(\bar{x} - \bar{y})$

$= (25 - 30) \pm 1.96 (1.56)$

$= -5 \pm 1.96(1.56)$

$= (-5 + 1.96(1.56) \; ; -5 - 1.96(1.56)$

$= (-1.94, -8.05)$ //

Scanned with CamScanner

Method-4

Test of Significance for Single Propertial for Large Samples.

Working Procedure:

Suppose a large random sample of size "n" has a sample propertial "p" that is propertial of success possessing a certain attribute.

To test the hypothesis that the propertion of in the population has a specified value P_0.

Then,

Null hypothesis ($H_0 = P = P_0$).

Alternate hypothesis $H_1 = P \neq P_0$ $[P < P_0$ (or) $P > P_0]$

Level of significance: α is given value.

Test statistics: $Z = \dfrac{P - p}{\sqrt{PQ/N}}$

where $P = \dfrac{x}{n}$, $Q = 1 - P$.

where "x" is the propertion of observed Given value.

Conclusion:

$Z_{cal} < Z_\alpha \rightarrow$ accept H_0

$Z_{cal} > Z_\alpha \longrightarrow$ Reject H_0.

Scanned with CamScanner

① In a sample of 1000 people, in Karnataka 540 are Rice Eaters and the rest are the wheat Eaters. Can we assume that both rice and wheat eaters equally popular yet at 1% of significance.

Sol: Given that:

$$n = 1000$$

$$P = 1/2, \quad q = 1/2$$

$$p = \frac{x}{n} \Rightarrow \frac{540}{1000} = 0.54$$

→ Null H (H_0): ($P = P_0$).

Both rice and wheat eaters are equally popular

i.e $P = P_0$

($P_0 = P$)

→ Alternate H (H_1): ($P \neq P_0$)

Both rice and wheat eaters are not equally popular

$P \neq P_0$. → Two tailed test.

→ Level of significance: (α)

$$\alpha = 1\%.$$

→ Test statistics:

$$z = \frac{p - P}{\sqrt{\frac{PQ}{N}}} = \frac{0.54 - 1/2}{\sqrt{\frac{1/2 \times 1/2}{1000}}}$$

$$\boxed{z = 2.52}$$

Scanned with CamScanner

Result :

$Z_{cal} = 2.52 \quad Z_\alpha = -2.58.$

$Z_{cal} < Z_\alpha.$

$2.52 < 2.58.$

∴ Accept the test H_0.

Both rice and wheat eaters are equally popular.

Q: Experience had shown that 20% of manufactured product of the top quality. In one day's production of 400 articles only 50 are the top quality. Test the hypothesis @ 0.05 level of significance.

Sol: Given that :

$n = 400 \qquad P = 20\% \rightarrow 20/100 = 0.2$

$x = 50 \qquad q = 0.8$

$$p = x/n = 50/400 = 1/8 = \boxed{0.125}$$

→ Null Hyp : (H_0) : $(P = P_0)$ $\quad (P_0 = p)$ small p

$P = P_0$

→ Alternate H (H_1) : $(P \neq P_0)$

$P \neq P_0$

→ Level of significance : (α)

$\alpha = 0.05 = 5\%.$

Scanned with CamScanner

iv) Test statistics :

$$z = \frac{P - P_0}{\sqrt{\frac{PQ}{n}}} = \frac{0.125 - 0.2}{\sqrt{\frac{0.2 \times 0.8}{400}}}$$

$$\boxed{Z = -3.75}$$

$$|Z|_{cal} = 3.75$$

v) Result :

$$|Z|_{cal} > Z_{tab(\alpha)}$$

$$3.75 > 1.96$$

Reject the test (H0) @ 5% level of significance

(5) In a random sample of 160 workes exposed to a certain amount of reduction 24 experienced sum ill effects construct a 99% confidence interval for the corresponding true percentage. 2%

Sol : Confidence limits (or) Confidence interval

$$= P \pm Z_{\alpha/2} \sqrt{PQ/N}$$

$n = 160$ $x = 24$ $P = \frac{x}{n} = 24/160 = 0.15$

$Q = 1 - P = 1 - 0.15 \Rightarrow 0.85$

$Z_\alpha = 2.58$ @ 99%

Scanned with CamScanner

confidence limits :

$$P \pm Z_\alpha \left(\sqrt{\frac{PQ}{N}}\right)$$

$$= 0.15 + 2.50\sqrt{\frac{0.15 \times 0.85}{100}},\ 0.15 - 2.50\sqrt{\frac{0.15 \times 0.8}{100}}$$

$$(0.22, 0.077)$$

$Z_\alpha = 0.25$ @ $\alpha = 99\%$.

Model No 5 :

Test of Equality for Two proportions :

Working Procedure :

Let P_1 and P_2 be the sample proportions in two large random samples of sizes n_1 and n_2, drawn from two population having P_1 and P_2.

To test wheather the two samples have drawn from same population.

1. $H_0 : P_1 = P_2$
2) $H_1 : P_1 \neq P_2$
3) Level of significance α is given

Scanned with CamScanner

Scanned with CamScanner

4)
a) When population properties known $z = \frac{P_1 - P_2}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}}$

b) When population properties are not known, but sample properties are known.

* Method of substitution: $z = \frac{p_1 - p_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$

* Method of Pooling: $z = \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

when $p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$

Conclusion:

$|z|_{cal} < z_\alpha \rightarrow$ Accepted (H_0)

$|z|_{cal} > z_\alpha \rightarrow$ Rejected H_0

① In two large population 30% and 25% resp. Pain healed people is the difference likely to be hidden in samples of 1200 & 900 resp. from the 2 population.

Given that :

Population properties are given $P_1 = 30\%$, $P_2 = 25\%$

$n_1 = 1200 \quad n_2 = 900$

$P_1 = 30\% = 30/100 = 0.3$

$P_2 = 25\% = 25/100 = 0.25$

Null H : $H_0 : P_1 = P_2$

There is no difference likely to be hidden sampling.

2) Alternate H :

There is difference likely to be hidden sampling

) level of Significance :

$\alpha = 5\%$ Assumed.

$Z\alpha = 1.96$.

) T.S :

$$\frac{P_1 - P_2}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}} = \frac{0.3 - 0.25}{\sqrt{\frac{0.3 \times 0.7}{1200} + \frac{0.25 \times 0.75}{900}}}$$

$P_1 = 0.3 \rightarrow Q_1 = 1 - P_1 = 1 - 0.3 = 0.7$ $\boxed{Z_{cal} = 2.65}$

$P_2 = 0.25 \rightarrow 1 - 0.25 = 0.75$

Scanned with CamScanner

⑤ Result:

$Z_{cal} > Z_\alpha$

$2.55 > 1.96$

Reject H_0.

② Among the items produced by a factory out of 800, 65 are defective and in other sample out of 300, 40 are defective. Test the significance diff b/w two proportions.

Sol: $n_1 = 800 > 30$; $n_2 = 300 > 30$, [Both are large samp

$x_1 = 65$, $x_2 = 40$

$P_1 = x_1/n_1 = 65/800 = 0.081$

$P_2 = x_2/n_2 = 40/300 = 0.133$

1) Null H_0: $H_0: P_1 = P_2$

There is no significance diff b/w sample proportion

2) Alternate H: $H_1: P_1 \neq P_2$

There is significance diff b/w sample proporti

3) L.O.S: $\alpha = 1\%$ @ $Z_\alpha = 2.58$

4. T.S:

$$Z = \frac{P_1 - P_2}{\sqrt{P\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

(By method of proby

Scanned with CamScanner

$$P = \frac{P_1 n_1 + P_2 n_2}{n_1 + n_2}$$

$$= \frac{0.081 \times 800 + 0.133 \times 300}{800 + 300}$$

$$\boxed{P = 0.095}$$

$$q = 1 - P = 1 - 0.095 = 0.905$$

$$\boxed{q = 0.905}$$

$$z \Rightarrow \frac{0.081 - 0.133}{\sqrt{0.095 \times 0.905 \left(\frac{1}{800} + \frac{1}{300}\right)}}$$

$$\underline{\underline{z = -2.61}}$$

$$|z|_{cal} = |-2.61| \quad z_{cal} = \underline{2.61}$$

5) Result:

$|z|_{cal} > z_\alpha \rightarrow$ Reject H_0

$2.61 > 2.58$

There is significance diff b/w Sample proportion

Scanned with CamScanner

③ A machine produced 20 defective articles in a batch 400 after over having it produced 10 defective in a batch of 300. Test @ 5% level, whether the machine improved.

Sol: $n_1 = 400$ $x_1 = 20$ $\alpha = 5\%$

$n_2 = 300$ $x_2 = 10$

$P_1 = x_1/n_1 = 20/400 = 0.05$

$P_2 = x_2/n_2 = 10/300 = 0.03$

Scanned with CamScanner

ulation.

Small samples :

For small sample size $n<30$, when σ, μ are unknown

we discuss the following distributions.

i) student 't' distribution.

ii) 'f' distribution.

iii) χ^2 [chisquare] distribution.

Degree of freedom :

The no. of independent variable which make up the statistic is known as degree of freedom and it is denoted by V.

→ The number of decrease of freedom is equal to the total number of observation less than the no. of independent variables.

Ex: In a set of

if k, no. of independent variables (or) 1

$$\boxed{V = n-k}$$

In general $\boxed{V = n-1}$

Scanned with CamScanner

1) 't' distribution:

If $x_1, x_2 \ldots x_n$ be any random sample of size drawn from a normal population with mean μ a variance σ^2, then the test statistic. $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$

where $s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$.

with $v = n-1$.

Procedure for t- distribution:

* The maximum possible probability with which we acc or reject a given statement is known as Level of signi
  → It is denoted by α. and it is also known as Siz
* We usually consider $\alpha = 5\%$.
* For a t-distribution α is considerd as $\frac{\alpha}{2}$

Problem 1:

A random sample of size 25 from a normal populati has mean $\bar{x} = 47.5$ and s.d (σ) is $= 8.4$. Does this information tends to support (or) refuse, the claim that mean of the population is 42.5.

Scanned with CamScanner

$25 < 30$ (small sample)

Given that :

The size of sample $(n) = 25$

$\bar{x}$ = Mean of sample = 47.5

μ = population mean = 42.5

(s) = 8.4.

[W]e have t- distribution $t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$

$= \frac{47.5 - 42.5}{\frac{8.4}{\sqrt{25}}}$

$t = 2.97 \simeq 2.98$ //.

[The] value of the t has 24 degree of freedom. i.e

$V = n - 1$

$= 25 - 1$

$\boxed{V = 24}$

$\alpha = 5\%$ i.e, 0.05

But in t distribution, $\alpha = \alpha/2$

$= \frac{0.05}{2} \rightarrow 0.025$

$t_{cal} = 2.98$

@ $v = 24$, if $\alpha = 0.025$

$t_{tab} = 2.064$ from the Table

$t_{cal} > t_{tab}$, $\therefore$ Reject the claim.

Scanned with CamScanner

3) A process for making ball bearings is under c[ontrol]
if the diametres of the bearing have a mean of 0.[5000 cm]
If a random sample of 10 of these bearings has
mean diameter. of 0.5060 cm and standard devia[tion]
of 0.0040 cm. is the process undu contsol or not.

Sol: Given:

n = size of sample = 10

$\bar{x}$ = Sample Mean = 0.5060

$\mu = 0.5000$

$S = 0.0040$

We have t - distsibution: $t = \dfrac{\bar{x} - \mu}{S/\sqrt{n}}$

$$= \frac{0.5060 - 0.5000}{\frac{0.0040}{\sqrt{10}}}$$

$$\boxed{t = 4.743}$$

where $v = n-1 = 10-1$ $\boxed{v=9}$

Let $\alpha = 0.05$ then for t distsibution $\alpha = \dfrac{\alpha}{2}$

$$= \frac{0.05}{2}$$

$$\boxed{\alpha = 0.025}$$

Scanned with CamScanner

$t_{tab} = 2.262$ [at $\alpha = 0.025$ and $v = 9$) from table

$t_{cal} > t_{tab}$.

$4.473 > 2.262$.

Rejected.

$\therefore$ The process is not under control

10 bearings made by the certain process have a mean diameter of 0.5060 cm. with sd of 0.004 cm. Assume that the data may be taken as a random sample from a normal distribution. Construct 95% confidence interval for the actual avg. diameter of variance.

Given:

Sol:

$n = 10$

$\bar{x} = 0.5060$.

$S = 0.004$

$\alpha = 5\% = 0.005$

$\alpha = \alpha/2 = \frac{0.005}{2} \Rightarrow 0.025$

Confidence limits:

$\Rightarrow \bar{x} \pm t_{\alpha/2} \left[\frac{s}{\sqrt{n}}\right]$

where $v = n-1$ i.e $10-1$ $\boxed{v=9}$

Scanned with CamScanner

Scanned with CamScanner

From table t_α value is t_{tab} i.e, 2.262

$\Rightarrow$ $0.5060 \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$

$$= 0.5060 \pm 2.262 \left[\frac{0.004}{\sqrt{10}}\right]$$

$$\Rightarrow 0.5060 + 2.262\left[\frac{0.004}{\sqrt{10}}\right], \quad 0.5060 - 2.262\left[\frac{0.004}{\sqrt{10}}\right]$$

$$(0.5031,\ 0.5089)$$

4) A sample of size 10 was taken from a popula[tion] s.d of sample 0.03 find the maximum error o[f] with 99% confidence.

Sol: Maximum Error formula $(E) = t_{\frac{\alpha}{2}}\left[s/\sqrt{n}\right]$

Given data:

$s = 0.03$

$n = 10$

$\alpha = 1\% = 0.01$

$\alpha = \frac{\alpha}{2} = \frac{0.01}{2} = 0.005$

$v = n-1 = 10-1 \Rightarrow \boxed{v=9}$

from table t_{tab}.

3.250

$$E = 3.250\left[\frac{0.03}{\sqrt{10}}\right]$$

$\therefore$ Maximum Error $\boxed{E = 0.030}$

Scanned with CamScanner

∴ confidence limits for 99% is 0.030.

find i) $t = 0.05$; when $v = 16$

ii) $t = 0.01$, when $v = 10$

iii) $t = 0.995$, when $v = 7$

i) Given:

$t = 0.05$

$v = 16$

from table $\Rightarrow t_{0.05} \Rightarrow 1.746$

ii) when $v = 10$, $t_{0.01}$

$= -2.764$ (we know that $t(\alpha) = -t(\alpha)$)

iii) when $v = 7$, $t_{0.995}$

$\Rightarrow t_{1-\alpha} = t_{1} - 0.095 \Rightarrow t_{0}$

$= 3.499$

Scanned with CamScanner

II) f - distribution:

An another important conditions, probability distrib. which plays an important rule in the sampling distribution is called f distribution.

→ It is used to determine whether the two sampl come from two populations have equal variances.

$$f = \frac{\text{Greater variance}}{\text{Smaller variance}} = \frac{S_1^2}{S_2^2}$$

where $(S_1^2 > S_2^2)$ or $(S_2^2 > S_1^2)$

$$S_1^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n_1 - 1}$$

$$S_2^2 = \sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{n_1 - 1} \quad \text{with } v_1 \text{ \& } v_2$$

③ for an f distribution find i) $f_{0.05}$ if $v_1 = 7, v_2 = 15$

ii) $f_{0.01}$, $v_1 = 24$, $v_2 = 19$

iii) $f_{0.95}$ with $v_1 = 19, v_2 = 24$

iv) $f_{0.99}$ with $v_1 = 28, v_2 = 1$

28

Scanned with CamScanner

Given:

$$f_{0.05}\underset{\;}{\overset{v_1,\,v_2}{[7, 15]}} = 2.71$$

) $f_{0.01}[24, 19] \Rightarrow \cancel{2.11}\ 2.92$

i) $f_{0.95}[19, 24] \Rightarrow$

No value in table so $f = 1 - 0.95 \Rightarrow \dfrac{1}{f_{0.05}(24, 19)}$

$$\boxed{f_\alpha = \frac{1}{f_\alpha(v_1, v_2)}}$$

$\dfrac{1}{2.11} \Rightarrow 0.473$

$f_{0.99}\overset{28,12}{(\cancel{28}, 12)}$

$\Rightarrow f_{1-0.01} = \dfrac{1}{f_{0.01}(v_1, v_2)}$

Here by reciprocal v_1 and v_2 values become inter...

$= \dfrac{1}{f_{0.01}(28, 12)} = \dfrac{1}{2.90}$

$= 0.344$

If two independent random samples of size $n_1 = 13$ and $n_2 = 7$ are taken from normal population. What is the probability that the variance of 1st sample will be atleast 4 times as large as their of the 2nd sample.

Sol: Given:

$n_1 = 13 \quad , \quad n_2 = 7$

Let s_1^2 and s_2^2 are variances.

$s_1^2 = 4 s_2^2$

But $s_1^2 > s_2^2$

$$F = \frac{\text{Greater.V}}{\text{Smaller.V}} = \frac{s_1^2}{s_2^2} = \frac{4\cancel{s_2^2}}{\cancel{s_2^2}}$$

$\boxed{F = 4}$ [F=4 in table to com

$V_1 = n_1 - 1 = 12 \quad (13 - 1)$

$V_2 = n_2 - 1 = 6 \quad (7 - 1)$

$\alpha = 0.05 \; (V_1, V_2) \quad 12, 6$

It is correct

(8) Come from normal population having same variance what is the probability that either of the sample variance will be atleast 7 times as large as the other. $n_4 = 8$.

Scanned with CamScanner

Given:

$n_1 = 8$ and $n_2 = 8$

Let s_1^2 and s_2^2 are variances

$$s_1^2 = 7 s_2^2$$

But $s_1^2 > s_2^2$

$$f = \frac{\text{Greater.V}}{\text{Smaller.V}} =$$

Scanned with CamScanner

15th χ^2 (chi)^2 distribution:

It is an continous probability distribution of a continous random variable x with probability density function $f(x) = \begin{cases} \frac{1}{2^{\nu/2}\,\Gamma \nu/2}\, x^{\nu/2-1} \cdot e^{-x/2} & \text{for } x>0 \\ 0, & \text{otherwise.} \end{cases}$

→ Sampling distribution of variance:

Let s^2 be the sample variance and σ^2 be the population variance then $\boxed{\chi^2 = \frac{(n-1)s^2}{\sigma^2}}$

with $\nu = n-1$ degree of freedom.

Problem.

① A manufactured claims the any of his list of items cannot have variance more than $1\,cm^2$. A sample of 25 items has a variance of $1.2\,cm^2$. Test whether the claim of manufacture is correct.

Sol: Given:

$n = 25$

population variance $\sigma^2 = 1\,cm^2$

Sample variance $s^2 = 1.2\,cm^2$

Scanned with CamScanner

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$= \frac{[25-1] \times 1.2}{1} = 28.8$$

$$\boxed{\chi^2 = 28.8}$$

At $\alpha = 5\% = 0.05$

$v = n-1 = 25-1$

$$\boxed{v = 24}$$

$\chi^2_{tab} = 36.415$ [at $\alpha = 0.05$ & $v = 24$ from table]

$\chi^2_{tab} > \chi^2_{cal}$

$\therefore$ Accept the claim.

The claim of manufactured is correct.

The claim that variance of a normal population is 21.3 is rejected. If the variance of the random sample of size 15 exceeds 39.74. What is the probability that the claim will be rejected even though $\sigma^2 = 21.3$.

Sol: Given

$n = 15$

$\sigma^2 = 21.3$

Scanned with CamScanner

$s^2 = 39.74$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$= \frac{(15-1) \times 39.74}{21.3}$$

$$\boxed{\chi^2 = 26.12}$$

$v = n - 1$

$15 - 1$

$\boxed{v = 14}$

$\chi^2_{tab} = 23.685$ at $v = 14$, $\alpha = 0.05$ from table

$\chi^2_{tab} < \chi^2_{cal}$

$23.685 < 26.12$

Reject claim.

The claim is rejected $\alpha = 0.05$ @ $v = 14$

@ $\chi^2 = 26.12$

③ They claim that the variance of normal population is $\sigma^2 = 4$ is to be rejected if the variance of random sample of size is 9

Scanned with CamScanner

exceeds 7.7535. What is the probability that the claim will be rejected.

Sol: Given:

$n = 9$

$\sigma^2 = 4$

$s^2 = 7.7535$

$$\psi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(9-1) \times 7.7535}{4}$$

$$\psi^2 = \underline{\underline{15.50}}$$

$v = n-1 = 9-1 \Rightarrow 8$

$\psi_{tab} = 15.507$

$\psi_{cal} = 15.507$

Rejected.

The claim is rejected when $\alpha = 0.05$ @ $\psi^2 = 15.50$

Scanned with CamScanner

Test of significance for small samples :

1) Students "t" test :

- An every important accept of the sampling theo is the study of test of significance which unable as to decide on the basic of the sample rejcts

* If the deviations b/w the observed sample stati and the hypothetical parameter value is significant.

* The deviation b/w two sample statistics is signific

Student t - test :

- Formula : Let $\bar{x}$ = Mean of a Sample

n = Size of Sample

σ = S.d of Sample

μ = Mean of Population supposed to normal

then students t - test defined by the sample

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}}$$

Scanned with CamScanner

If s^2 = sample variance then $s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$

$$\frac{S^2}{s^2} = \frac{n}{n-1} \text{ (or) } S = \sqrt{\frac{n}{n-1}} \cdot s$$

Student - t test for single Mean :

uppose we want to test a) if a random "x", of size "n" have been drawn from a normal population with a mean "μ"

b) if the sample mean diff gnificantly from the hypothesis value μ of the population mean.

In this case, the statistic is given by $t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n-1}}}$

roblems :

A sample of 26 bulbs gives a mean life of 990 hours with a s.d of 20 hours. The manufactured laims that the mean life of bulbs is 1000 hours s the sample is not up to the standard?

Sol: Given data :

sample size $(n) = 26 < 30$ (Small sample)

Scanned with CamScanner

$S = 20$

$\bar{x} = 990$

population mean $(\mu) = 1000$

Degree of freedom $(\nu) = n-1 = 26-1 \Rightarrow 25$.

Here W.K.T $\bar{x}, \mu, s$ and n (so) we use the student t

i) N. Hypothesis : $H_0 : \bar{x} = \mu$.

The sample is upto the standard

ii) A.H : $H_1 : \bar{x} \neq \mu$.

The sample is not upto the standard.

iii) level of significance :

$\alpha = 0.05 \Rightarrow \alpha = \frac{0.05}{2} = 0.0025$

iv) T.S :

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} = \frac{990 - 1000}{\frac{20}{\sqrt{26-1}}} \Rightarrow -2.5$$

$|z|_{cal} = |-2.5| \Rightarrow 2.5$

z_{tab} ~~1.960~~ 2.060

$z_{tab} < z_{cal} \rightarrow 2.5 > \cancel{1.968}$ 2.060 Reject

$\therefore$ The sample is not upto the standard

A random sample of 6 steel beams has a mean of strength of 58,392 [PSI] ~~pounds~~ with a standard deviation of 648 PSI. Use the information & level of significance $\alpha = 0.05$ to test whether the true average of strength of the steel from which this sample came is 58000 psi. Assume normally.

Sol:

Given that:

$$\mu = 58000$$

$$n = 6$$

$$\bar{x} = 58,392 \text{ Psi}$$

$$\alpha = 0.05$$

$$s = 648$$

degree of freedom $(\nu) = n-1 = 6-1 = 5$.

1) Null H: $H_0 : \bar{x} = \mu$.

2) A.H: $H_1 : \bar{x} \neq \mu$

3) $\alpha = \frac{0.05}{2} \Rightarrow$ ~~0.025~~ 0.0025

4) T.S: $t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = \frac{58392 - 58000}{\frac{648}{\sqrt{6-1}}} = 1.352$

Scanned with CamScanner

$t_{cal} = 1.352$

$t_{tab} = \cancel{2.015}\ 2.571$

$t_{cal} < t_{tab}$.

Accept the H_0.

Hence the avg strength of steel is equal to 580

Note:

If t-test we, have to take level of significance $\alpha = \alpha/2$ in two tailed test, otherwise α is taken

16/10

③ A random sample of 10 ~~po boys~~ boys had the follo

IQ's 70, 120, ~~110~~ 101, 88, 83, 95, 98, 107 and 100.

a) Does the data support the assumption of a populati

mean IQ of 100.

b) Find the reasonable range in which most of the

mean IQ values of samples of 10 boys lie.

Scanned with CamScanner

) Here, we have to find s.d and mean.

$$\text{Mean}(\bar{x}) = \frac{\Sigma x}{m} = \frac{972}{10} = 97.2.$$

x	$x-\bar{x}$	$(x-\bar{x})^2$
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	21.84
98	0.8	0.64
107	9.8	96.84
100	2.8	7.84
972		1833.60

$$\Sigma(x-\bar{x})^2 = 1833.60$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x-\bar{x})^2$$

Scanned with CamScanner

$$s^2 = \frac{1833.60}{10-1} = 203.73$$

$$s^2 = 203.73$$

$$s = \sqrt{203.73}$$

$$\boxed{s = 14.27}$$

i) N.H : $H_0 : \bar{x} = \mu$.

The data supports the assumption of population mean ~~IQ~~ of 100 in the population.

ii) A.H : $H_1 = \bar{x} \neq \mu$.

The data not supports.

iii) α : $0.05 \Rightarrow \frac{0.05}{2} = 0.025$ //

iv) T.S : $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} = \dfrac{97.2 - 100}{\frac{14.27}{\sqrt{10-1}}}$

$$\boxed{t = -0.588}$$

v) $v = n-1 \Rightarrow 10-1 \quad v = 9$ [Z] = 10.5

$t_{tab} = 2.262$

$t_{cal} = 0.588$

Scanned with CamScanner

$t_{tab} > t_{cal}$

Accept the H_0

∴ The data supports the assumption of popula[tion] mean IQ of 100 in the population

3) Reasonable ranges.

The 95% confidence limits are given by :-

$$\Rightarrow \bar{x} \pm t_\alpha \, s/\sqrt{n}$$

$$\left[\bar{x} \pm t_{0.025} \frac{s}{\sqrt{10}}\right]$$

$$= \left[97.2 + 2.262 \times \frac{14.27}{\sqrt{10}}, \; 97.2 - 2.262 \times \frac{14.27}{\sqrt{10}}\right]$$

$$[107.4, \; 87]$$

Q) 8 students were given a test in statistics and after 1 month coaching, they were given another test of the similar nature. The following table gives the increase first their marks in the second test over the first.

Student No.	1	2	3	4	5	6	7
Increase of marks	4	-2	6	-8	12	5	-7

Do the marks indicate that the students have gained [from] the coaching.

Sol: Mean $(\bar{x}) = \frac{\Sigma x}{N} = \frac{36}{8} = 4.5$

x	$(x-\bar{x})$	$(x-\bar{x})^2$
1	-3.5	12.25
2	-2.5	6.25
3	-1.5	2.25
4	-0.5	0.25
5	0.5	0.25
6	1.5	2.25
7	2.5	6.25
8	3.5	12.25
36		42.

$\Sigma(x-\bar{x})^2 = 42$

$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x-\bar{x})^2$

$s^2 = \frac{42}{8-1} = 6$

Scanned with CamScanner

Sol:

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{4-2+6-8+12+5-7+2}{8}$$

$$= 12/8$$

$$\boxed{\bar{x} = 1.5}$$

x_i	$(x-\bar{x})$	$(x-\bar{x})^2$
4	2.5	6.25
-2	-3.5	12.25
6	4.5	20.25
-8	-9.5	90.25
12	10.5	110.25
5	3.5	12.25
-7	-8.5	72.25
2	0.5	0.25
12		324

Scanned with CamScanner

$(\Sigma x - x^2) = 324$

$$s^2 = \frac{324}{8-1} = \frac{324}{7} = 46.28$$

$$s = \sqrt{46.28} = 6.80$$

$$\boxed{S = 6.80}$$

i) N.H: $H_0 : \bar{x} = \mu$.

The test provides no evidence that the students have been benefited by course

ii) A.H: $H_1 : \bar{x} \neq \mu$

iii) L.O.S: $\alpha = 0.05 \rightarrow \frac{0.05}{2} = 0.025$.

iv) T.S: $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} = \dfrac{1.5 - 0}{\frac{6.80}{\sqrt{8-1}}} \Rightarrow 0.58$

v) $t_{cal} = 0.58 \Rightarrow |z| = |0.58|$

$t_{tab} = 2.365$

$t_{\alpha} = 0.58$

$t_{tab} > t_{\alpha}$

Accepted.

Scanned with CamScanner

Students t-Test for difference of means

Procedure :

Let $\bar{x}$ and $\bar{y}$ be the means of two independent samples of sizes n_1 and n_2 where $n_1 < 30$ & $n_2 < 30$. drawn from two populations having means μ_1 and μ_2. To test whether the two population means are equal.

Null Hyp: $\mu_1 = \mu_2$

Alt Hyp: $H_1 = \mu_1 \neq \mu_2$

If $\sigma_1 = \sigma_2 = \sigma$, then an unbaised estimate s^2 of the common variance σ^2 is given by

$$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}.$$

$$\text{Standard Error of } (\bar{x} - \bar{y}) = s\sqrt{1/n_1 + 1/n_2}.$$

$$\text{where } s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}.$$

iii) Test statistic $t = \dfrac{\bar{x} - \bar{y}}{s\sqrt{1/n_1 + 1/n_2}}$

where with $v = n_1 + n_2 - 2$.

Scanned with CamScanner

where $\bar{x} = \frac{\Sigma x_i}{n_1}$ and $\bar{y} = \frac{\Sigma y_i}{n_2}$.

$$S^2 = \frac{\Sigma(x_i - \bar{x})^2 + \Sigma(y_i - \bar{y})^2}{n_1 + n_2 - 2}.$$

v) Result : $|t|_{cal} < t_\alpha$.

$\hookrightarrow$ Accepted otherwise reject H_0

Note :

i) Confidence limits :

95% confidence limits for the difference of two population means are

$$(\bar{x} - \bar{y}) \pm t_\alpha S \sqrt{1/n_1 + 1/n_2}$$

where $\alpha = 0.025$

ii) 99% confidence limits for the difference of two population means are :

$$\bar{x} - \bar{y} \pm t_\alpha S \sqrt{1/n_1 + 1/n_2}$$

where $\alpha = 0.005$

1%

100 = 0.

0.0/2

= 5.

Scanned with CamScanner

Two independent samples of 8 and 7 items respectively had the following values.

ample - 1	11	11	13	11	15	9	12	14
ample - II	9	11	10	13	9	8	10	—

Is the difference b/w the means of the samples significant.

Sol: Given that :

$$n_1 = 8 \quad , \quad n_2 = 7$$

$$\bar{x} = \frac{\Sigma x_i}{N} = 96/8 = 12$$

$$\bar{y} = \frac{\Sigma y_i}{N} = 70/7 = 10.$$

x	y	$(x-\bar{x})^2$	$(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$
11	9	-1	-1	1	1
11	11	-1	1	1	1
13	10	1	0	1	0
11	13	-1	3	1	9
15	9	3	-1	9	1
9	8	-3	-2	9	4
12	10	0	0	0	0
14	—	2 / 26	—	4	— / 16

Scanned with CamScanner

$$t = \frac{\bar{x} - \bar{y}}{s\sqrt{1/n_1 + 1/n_2}}$$

$$s^2 = \frac{\Sigma (x_i - \bar{x})^2 + (y_i - \bar{y})^2}{n_1 + n_2 - 2}$$

$$= \frac{26 + 16}{8 + 7 - 2}$$

$$\boxed{s^2 = 3.230} \qquad s = \sqrt{3.230} = 1.0$$

$$t = \frac{12 - 10}{1.79\sqrt{1/8 + 1/7}}$$

$$\boxed{t = 2.15}$$

$$V = n_1 + n_2 - 2$$

$$8 + 7 - 2$$

$$\boxed{V = 13}$$

$t_{tab} = 2.160$ @ $\alpha = 0.025$

$t_{cal} < t_{tab}$

Accepted H_0.

Scanned with CamScanner

Paired Sample t-test

(x_1,y_1), (x_2,y_2) ---- x_n, y_n be the pairs of x_1 and y_1, x_2 and y_2 ---- x_n, y_n be the pairs of before and after the propertion in the problem applied paired t-test to examine the significance of difference of two solutions.

$$d_i = x_i - y_i$$

where $i = 1,2,3 \text{ ------ } n.$

Null H: $\mu_1 = \mu_2$

There is no significant difference b/w the means in two solutions.

Alt. H: $\mu_1 \neq \mu_2$

α :- given

$t = \frac{\bar{d} - \mu}{s/\sqrt{n}} = \bar{d}/s/\sqrt{n}$ where $(\mu = 0)$

where $\bar{d} = \frac{\Sigma d_i}{n}$

$s^2 = \frac{1}{n-1} \sum_{i=1} (d_i - \bar{d})^2$

Result:

If $t_{cal} < t_{tab}(v)$ Accepted

(or) $s^2 = \frac{\Sigma d_i^2 - n(\bar{d})^2}{n-1}$

where v = degree of freedom

$v = n-1$

Scanned with CamScanner

Q) Memory capacity of 10 students were tested before after training. State whether the training was effecti or not from the following scores.

Before Training	12	14	11	8	7	10	3	0	5
After Training	15	16	10	7	5	12	10	2	3

Sol: Let $H_0 = \mu_1 = \mu_2$.

There is no significant effect of Training.

→ A.H: $H_1 = \mu_1 \neq \mu_2$.

There is significant effect.

→ $\alpha = \frac{0.05}{2} = 0.025$

→ Calculation for $\bar{d}$ and s:

Memory of capacity

Before training (x)	After training (y)	$d = x - y$	d^2
12	15	-3	9
14	16	-2	4
11	10	1	1
8	7	1	1
7	5	2	4

Scanned with CamScanner

	12	8 -2	64
10	10	-7	49
3	2	-2	24
0	3	2	4
5	8	-2	4
6			
		-12	84

$\Sigma d = -12$, $\Sigma d^2 = 84$

$$\bar{d} = \Sigma d/n = -12/10 = -1.2$$

$$s^2 = \frac{\Sigma d^2 - n(\bar{d})^2}{n-1}$$

$$= \frac{84 - 10(-1.2)^2}{10-1}$$

$$s^2 = 7.73$$

$$s = \sqrt{7.73}$$

$$\boxed{s = 2.78}$$

$$t_{cal} = \frac{\bar{d} - 0}{\frac{s}{\sqrt{n}}} = \frac{-1.2 - 0}{\frac{2.78}{\sqrt{10}}} = -1.36$$

$$|z| = |-1.36| = 1.36$$

Scanned with CamScanner

$v = n-1 \Rightarrow 10-1 \Rightarrow 9.$

$\alpha = 0.025$

$t_{cal} = 1.365$, $t_{tab} = 2.262$

$t_{cal} < t_{tab}$

Accepted.

Imp

→ The blood pressure of 5 women before and after in take of a certain drug are given below.

Before	110	120	125	132	125
After	120	118	125	136	121

Test whether the significance of drug 1% level of significance.

Scanned with CamScanner

Snedecor's F-test of Significance :

Procedure :

When testing of significance of the difference of means of two samples, we assumed that two samples came from the same population are from populations with equal variances.

→ If the variances of the population are not equal a significant difference means may be arranged. Hence we before we apply t-test for the significance of difference of two means, we have to test for the equality of population variances using F-test of significance.

If s_1^2 and s_2^2 are the variance of two samples sizes n_1 and n_2 respectively then the population variances are given by $n_1 s_1^2 + n_2 s_2^2 = (n_2 - 1) S_2^2$.

$$n_1 s_1^2 = (n_1 - 1) S_1^2 + n_2 s_2^2 = (n_2 - 1) S_2^2$$

The Quantities $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ are call degrees of freedom of these estimate

S_1^2 and S_2^2 are significantly different (or) if the samples may be drawn from a same populatio

Scanned with CamScanner

(or) from two populations with same variance σ^2.

Test for Equality of two population variances :

Procedure :

Let two independent random samples of sizes n_1 and n_2 be drawn from two normal populations.

To test hypothesis; that the two population variances σ_1^2 and σ_2^2 are equal.

→ Let the null hypothesis (NH) = ~~H_0~~ $\sigma_1^2 = \sigma_2^2$.

→ Alternate Hypothesis (AH) : $\sigma_1^2 \neq \sigma_2^2$ → α i.e. = Given value

The estimates of σ_1^2 and σ_2^2 are given by :

$$S_1^2 = \frac{n_1 s_1^2}{n_1 - 1} = \frac{\Sigma (x_i - \bar{x})^2}{n_1 - 1}$$

$$S_2^2 = \frac{n_2 s_2^2}{n_2 - 1} = \frac{\Sigma (y_i - \bar{y})^2}{n_2 - 1}$$

where s_1^2 and s_2^2 are variances of the two samples.

→ Test statistic :

$$F = \frac{S_1^2}{S_2^2} \text{ (or) } S_2^2 / S_1^2$$

where $S_1^2 > S_2^2$ (or) $S_2^2 > S_1^2$

i.e. F distribution with $(n_1 - 1, n_2 - 1)$ are degree of freedom

Scanned with CamScanner

Conclusion:

If $f_{cal} < f_{tab} \rightarrow$ Accept H_0

$f_{cal} > f_{tab} \rightarrow$ Reject H_0.

Problems:

① The nicotants contains in milligrams in 2 samples tobacco where found to be follows:

Sample-A	24	27	26	21	25	-
Sample B	27	30	28	31	22	36.

Can it be set that the two samples have come from the normal population. Ⓑ Student t test

Sol: → To test whether the two samples came from same normal population.

Sol: Given that: $n_1 = 5$ and $n_2 = 6$

x	$x-\bar{x}$	$(x-\bar{x})^2$	y	$(y-\bar{y})$	$(y-\bar{y})^2$
24	-0.6	0.36	27	-2	4
27	-2.4	5.76	30	-1	1
26	1.4	1.96	28	-1	1
21	-3.6	12.96	31	2	4
25	+0.4	0.16	22	-7	49
-	-		36	7	49
123		21.2	174		108

Scanned with CamScanner

$\bar{x} = \Sigma x_i / N = 123/5 = 24.6$

$\bar{y} = \Sigma y_i / N = 174/6 = 29.$

$$S_1^2 = \frac{\Sigma (x_i - \bar{x})^2}{n_1 - 1}$$

$$= \frac{21.2}{5-1} = \frac{21.2}{4} = 5.3$$

$$S_2^2 = \frac{\Sigma (y_i - \bar{y})^2}{n_2 - 1} = 108/6-1 = 108/5 = 21.6$$

$f = S_2^2 / S_1^2 = 21.6 / 5.3 \Rightarrow 4.07$

$f_{cal} = 4.075$

N.H: $H_0 = \sigma_1^2 = \sigma_2^2$

2) A.H: $H_1 = \sigma_1^2 \neq \sigma_2^2$

3) $\alpha = 0.05$

4) $f_{cal} = 4.07$

5) f_{tab} where $v_1 = n-1 \Rightarrow 5-1 \Rightarrow 4$

$v_2 = 6-1 = 5$

$f_{tab} = 5.19$

$f_{tab} > f_{cal}$

Accepted $H_0 = \sigma_1^2 = \sigma_2^2$

Scanned with CamScanner

(b) Student t - test :

Two means are equal $\mu_1 = \mu_2$

Given $\bar{x} = 24.6$ and $\bar{y} = 29$.

$\Sigma(x_i - \bar{x})^2 = 21.2$

$\Sigma(y_i - \bar{y})^2 = 108$

$$S^2 = \frac{1}{n_1+n_2-2}\left[\Sigma(x_i-\bar{x})^2 + \Sigma(y_i-\bar{y})^2\right]$$

$$S^2 = \frac{1}{5+6-2}\left(21.2+108\right)$$

$$S^2 = 14.35$$

$$S = \sqrt{14.35} \rightarrow \boxed{S = 3.78}$$

$$t = \frac{\bar{x} - \bar{y}}{S\sqrt{1/n_1 + 1/n_2}}$$

$$t = \frac{24.6 - 29}{\left[3.78\sqrt{1/5 + 1/6}\right]}$$

$$t = -1.92$$

$$|t| = \underline{\underline{1.92}}$$

$v = n_1 + n_2 - 2$

$5 + 6 - 2$

$\boxed{v = 9}$

$\alpha = \frac{0.05}{2}$

$\alpha = 0.0025$

t_{tab} @ $\alpha = (0.025, 9) \rightarrow 2.262$

$t_{tab} > t_{cal}$

$\therefore H_0$ Accepted

Q) In one sample of 10 observations, the sum of sequences of the deviations of the sample values from sample mean was 120. and other sample of 12 observat it was 314. Test whether the difference is significant at 5% level.

Sol:

N.H: (H_0): $\sigma_1^2 = \sigma_2^2$ {where σ_1^2 & σ_2^2 are the variances of the two populations from samples are drawn}

A.H: (H_1): $\sigma_1^2 \neq \sigma_2^2$

α: 5% = 0.005

T.S: $f = s_2^2 / s_1^2$

$$s_1^2 = \frac{\Sigma(x_i - \bar{x})^2}{n_1 - 1} = \frac{120}{10-1} = 120/9$$

$= 13.33$

Here $n_1 = 10$, $n_2 = 12$

Scanned with CamScanner

$$S_2^2 = \frac{\Sigma(y_1 - \bar{y})^2}{n_2}$$

$$= \frac{314}{12-1} = \underline{\underline{28.54}}.$$

$$f \Rightarrow \frac{S_2^2}{S_1^2}$$

$$\frac{28.54}{13.33} = \underline{\underline{2.14}}$$

$f_{cal} = 2.14$

Result:

$f_{tab} = 4.63$ $v_1 = 9$, $v_2 = 11$ @ 0.05 //

$f_{tab} > f_{cal}$

Accepted H_0

(6)

Scanned with CamScanner