UNIT – IV

RANDOM VARIABLES & DISTRIBUTIONS

Discrete & continuous

Distributions: Bionomial distribution, poisson distribution and their properties, Normal distribution, sampling distribution of mean(σ is known & Unknown).

Factorial notation: The product of n consecutive +ve integers denoted by $n!$ (or) $\underline{|n}$ $= n\times(n-1)\times(n-2)\times \text{------} 3\times2\times1$

Ex: $6! = 6\times5\times4\times3\times2\times1 = 720$

$0! = 1$

Permutation (nP_r): the no. of permutations of n different things taken 'r' at a time. Here, repetations are allowed.

$$nP_r = n(n-1)(n-2) \text{-----} n-(r+1) \text{ i.e}$$

$$\text{i.e } nP_r = \frac{n!}{(n-r)!} \qquad 3P_2 = \frac{3!}{(3-2)!} = 6$$

$$nP_n = n!$$

$$nP_0 = 1$$

Combinations (nC_r): the no. of combinations of n different things taken 'r' at a time.

$$nC_r = \frac{n(n-1)(n-2) \text{------} \overset{r \text{ factors}}{\cancel{n-(r+1)}}}{r!}$$

$$\text{i.e } nC_r = \frac{n!}{r!(n-r)!}$$

$$nC_n = 1$$

$$nC_0 = 1$$

$$3C_2 = \frac{3!}{2!(3-2)!} = \frac{6}{2} = 3$$

Ex: All permutations made with the letters a,b,c by taking 2 at a time on a,b: ab, bc, ca, ba, cb, ac

$||^{ly}$ All the combinations formed by 2 at a time on ab, bc, ca

Experiment: An operation which can produce some well defined outcomes is called an Experiment (or) physical action (or) process i.e observed and the result is noted.

Ex: when we toss a coin then either head (or) tail appears.

Firing a missile

Determinate (st)

Deterministic (or) predictable experiment: If the results of an experiment is unique then the experiment is said to be predictable experiment

Ex: If r is the radius of a circle, then its area is πr^2 which gives uniquely the area of circle.

Probabilistic (or) unpredictable experiment:
If the results of an experiment is not unique (or) not certain but may be one of the several possible outcomes then the experiment is said to be probabilistic (or) unpredictable experiment.

Ex: By looking at the sky, one is not sure if it rains (or) not

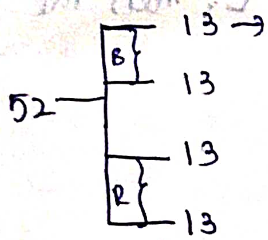
Random experiment: If an experiment is conducted any number of times and essentially identical conditions, there is a set of all possible outcomes associated with it. If the result is not certain and is any one of the several possible outcomes than the experiment is called Random experiment (or) Random trial.

The outcomes are called Elementary events and set of outcomes is an event.

Ex: 1. Tossing a fair coin

2. Rolling an unbiased die

3. Drawing a card from a pack of well shuffled cards



4. Picking up a ball of certain colour from a bag containing balls of different colour.

Sample space: The set of all possible outcomes of a random experiment is called sample space.

Ex: In tossing a coin, the possible outcomes are H, T

$S = \{H, T\}$

If two coins are tossed at a time then the set of possible outcomes are $S = \{HH, TT, HT, TH\}$

If a rolling a die, the possible outcomes are $S = \{1, 2, 3, 4, 5, 6\}$

If a coin tossed n times, n coins are tossed at a time, then the no. of possible outcomes are $S = 2^n$.

Event: An event in a trial that cannot be further split is called a simple event (or) an elementary event (or) Any subset of a sample space is called an event.

Ex: If a die is thrown then the possible outcomes $S = \{1, 2, 3, 4, 5, 6\}$

Getting an even no. is an event.

Types: 1. Equally likely event

2. Exhaustive event

3. Mutually exclusive events

Probability: probability is the branch of mathematics which deals the influence 'of chance' is the theory of probability.

It is a concept which numerically measures the degree of certainity, uncertainity (or) occurence (or) non occurence of an event.

It is a measurement of uncertainity (or) chances of the happening of events are considered.

If an experiment is conducted, let n is the total no. of outcomes. m is the favourable cases of an event E, Then the probability of occurence of an event E is

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

Axioms (or) properties of probability:

Let S be a finite sample space. If E is an event of sample space S then $0 \leq P(E) \leq 1$

ii. $P(S) = 1$

iii. $P(\phi) = 0$

iv. $E_1 \subset E_2 \Rightarrow P(E_1) \leq P(E_2)$

v. E_1 & E_2 are disjoint events then probability of $P(E_1 \cup E_2) = P(E_1) + P(E_2)$

vi. For any events E_1 & E_2 $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

vii. If $E_1, E_2, \ldots\ldots E_k$ are k mutually exclusive events in the sample space of an experiment then $\sum_{i=1}^{k} P(E_i) = 1$

viii. If $\bar{E}$ denotes not E then $P(\bar{E}) = 1 - P(E)$ $\quad 0 \leq P(E) \leq 1$

$0 \leq P(\bar{E}) \leq 1$

1. A class consists of 6 girls and 10 boys. If a committee of 3 is choosen at random, from the class, Find the probability that

i. 3 boys are selected ii. Exactly 2 girls are selected

Total no. of students = 16

i) The total no. of ways of choosing 3 students out of 16

$n(S) = 16C_3 = \frac{16!}{3!(16-3)!}$ $\qquad nC_r = \frac{n!}{r!(n-r)!}$

Let E be the event of selecting boys 3 out of 10, this can be done as $n(m) = 10C_3$

The probability of selecting 3 boys is $P(E) = \frac{n(m)}{n(S)}$

$$= \frac{\frac{n!}{r!(n-r)!}}{\frac{n!}{r!(n-r)!}} = \frac{\frac{10!}{3!(10-3)!}}{\frac{16!}{3!(16-3)!}}$$

$$= \frac{3}{14}$$

ii) Let E be the event of selecting 2 girls out of 6 this can be done as $6C_2$

∴ No. of favourable outcomes for event E is $n(m) = 6C_2 \cdot 10C_1$

The probability of selecting exactly 2 girls is $P(E) = \frac{n(m)}{n(S)}$

$$= \frac{6C_2 \cdot 10C_1}{16C_3}$$

$$= \frac{\frac{6!}{2!(6-2)!} \cdot \frac{10!}{1!(10-1)!}}{\frac{16!}{3!(16-3)!}} = \frac{150}{560} = \frac{15}{56}$$

2. A pair of dice are thrown, find the probability of the event in which the sum is 10 or more.

Let n(S) be the no. of outcomes when a pair of dice are thrown $n(S) = 6^2 = 36 = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6)$

$(2,1), (2,2), (2,3), (2,4), (2,5), (2,6)$

⋮

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

Let E be the event of getting sum is 10 (or) more than 10 is

$n(E) = \{(4,6), (5,5), (5,6), (6,4), (6,5), (6,6)\}$
$= 6$

The probability of the event getting the sum is 10 (or) more is $P(E) = \frac{n(E)}{n(S)} = \frac{6}{36} = \frac{1}{6}$

Random variable: A real variable 'x' whose value is determined by the outcome of a random experiment is called a Random variable.

Random variable is a real valued function which maps the numerical (or) Non-numerical sample space (S is the domain & R is the range) of the random experiment to the real values (codomain (or) range) i.e $X : S \rightarrow R$

Ex: Consider a random experiment consisting of tossing a coin twice. The sample space $S = \{s_1, s_2, s_3, s_4\}$

where $s_1 = HH$ $s_2 = HT$ $s_3 = TH$ $s_4 = TT$

Define a function $X : S \rightarrow R$ by $X(s)$ = no. of heads, then

$X(s_1) = HH = 2$

$X(s_2) = HT = 1$

$X(s_3) = TH = 1$

$X(s_4) = TT = 0$

Range of $(X) = [X(s), s \in S]$

$= \{X(s_1), X(s_2), X(s_3), X(s_4)\}$

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

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

Types of Random variables (RV): 1. Discrete Random variables (DRV)

A Random variable X which can take only a finite no. of discrete values in an interval of domain is called discrete Random variable.

(or)

If the random variable take the values only on the set $\{0, 1, 2, 3, \ldots\ldots n\}$ is called DRV.

The above ex. is a DRV.

The no. of defectives in a sample of electric bulbs.

The no. of printing mistakes in each page of a book.

The no. of telephone calls received by the telephone operator.

2. continuous Random variables (CRV): A RV which can take values continuously which takes all possible values in a given interval is called a CRV.

Ex: Height, Age, weight, Temperature ... etc which are changing continuously.

probability Distribution function (PDF): Let x be a RV then the PDF associated with x is $F_X(x) = F(x) = P(X \le x) =$

$P[s : X(s) \le x\}$

$-\infty \le x \le \infty$

Discrete probability distribution function: suppose x is a RV with possible outcomes are $x_1, x_2 - - - - - x_n$. The probability of each possible outcome x_i is $P(X_i) = P(X = x_i)$

$= P(x_i) \quad i = 1, 2, - - - - n$

If $P(x_i)$ satisfies the two conditions i. $P(x_i) \ge 0$ $\quad i = 1, 2, - - - n$

ii. $\sum P(x_i) = 1$

then the function P is called the probability mass function 5)

PMF) of RV 'x'. And a set of $P(x_i)$ is called the (DPDF) of RV 'x'.

The PDF of RV 'x' is

'x' is	x_1,	x_2, - - - - - -	x_n
$P(X=x_i)$	P_1	P_2 - - - - - -	P_n

and

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

Probability of getting no heads $P(X=0) = P(TT) = \frac{1}{4}$

probability of getting 1 head

$P(X=1) = P\{(HT), (TH)\} = \frac{2}{4} = \frac{1}{2}$

probability of getting 2 heads

$P(X=2) = P[HH\} = \frac{1}{4}$

$X = x_i$	0	1	2
$P(X = x_i)$	1/4	1/2	1/4

Expectation: A RV x assumes the values $x_1, x_2, \ldots, x_n$ w.r.t the probabilities of $P_1, P_2, \ldots, P_n$ then the expectation are mean(or) expected value of x is denoted by

$$E(x) = \text{mean}(\mu) = \sum_{i=1}^{n} P_i x_i$$

$$= P_1x_1 + P_2x_2 + \ldots + P_nx_n$$

$$E\{g(x)\} = \sum P_i\{g(x_i)\}$$

$$E(X+Y) = E(X) + E(Y)$$

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

$$E(X.Y) = E(X).E(Y)$$

$$E(K) = K$$

$$E(KX) = KE(X)$$

If x is a RV and a, b are constants then the expectation of $E(ax+b) = aE(X) + b$

$$E(X+Y+Z) = E(X) + E(Y) + E(Z)$$

$$E(aX+bY) = aE(X) + bE(Y)$$

$$E(X - \bar{X}) = 0$$

$$E(XYZ) = E(X).E(Y).E(Z)$$

The mean μ the mean value μ of the Discrete distribution function is given by $\mu = \frac{\sum P_i x_i}{\sum P_i = 1} = (\sum P_i x_i)$

Variance (σ^2): The variance of the DPDF of a RV x is the mathematical expectation and is denoted by $\sigma^2 = E(x-\mu)^2$

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

$$= P_i x_i^2 - \mu^2$$

Standard deviation (σ):

$$\sigma = \sqrt{P_i x_i^2 - \mu^2}$$

If x is DRV and a, b are constants then $V(ax+b) = a^2 V(x)$

where $V(x)$ is variance of x

If $a=1$ then $PV(ax+b) = V(a x+b) = V(x)$

If $a=0$ then $V(b) = 0$

1. A RV x has following probability function.

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

Determine a) k b) evaluate $P(X<6)$, $P(X \geq 6)$, $P(0<X<5)$ & $P(0 \leq X \leq 4)$. c) If $P(X \leq k) > \frac{1}{2}$ find the minimum values of k. d) Determine the Distribution function of X. e) mean f) variance & standard deviation.

a) We know that sum of probability = 1

$$\sum_{i=1}^{7} P x_i = 1$$

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

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

$$10k^2+10k-9k \Rightarrow 10k^2+10k-9k-1=0$$

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

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

$$k = \frac{1}{10}, -1$$

$k = \frac{1}{10}$ ($\because P(x_i) \geq 0$)

b) $\sum_{i=0}^{5} P(x_i)$

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

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

$= k^2+8k$

$\left(\frac{1}{10}\right)^2 + 8\left(\frac{1}{10}\right) = 0.81$

$P(X \geq 6) = 1 - P(X<6)$ (or) $\sum_{i=6}^{7} P(x_i)$

$= 1 - 0.81$ (or) $P(X=6) + P(X=7)$

$= 0.19$ $\quad 2k^2+7k^2+k$

$= 9k^2+k$

$= 9\left(\frac{1}{10}\right)^2 + \left(\frac{1}{10}\right)$

$= 0.19$

$P(0<X<5) = \sum_{i=1}^{4} P(x_i)$

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

$= k + 2k + 2k + 3k = 8k = 8\left(\frac{1}{10}\right) = \frac{4}{5} = 0.8$

$P(0 < x \leq 4) = \sum_{i=0}^{4} P(x_i)$

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

$= 0 + k + 2k + 2k + 3k = 8k = 0.8$

c) $P(X \leq k) > \frac{1}{2}$

The required minimum value of k is obtained as

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

$= 0 + k = \frac{1}{10} = 0.1 \ngtr \frac{1}{2}$

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

$= 0 + k + 2k = 3k = 3\left(\frac{1}{10}\right) = 0.3 \ngtr \frac{1}{2}$

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

$= 0 + k + 2k + 2k = 5k = 5\left(\frac{1}{10}\right) = 0.5 \ngtr 0.5 = \frac{1}{2}$

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

$= 0 + k + 2k + 2k + 3k = 8k = 0.8 > 0.5 = \frac{1}{2}$

The minimum value of k is 0.8

d) The distribution function of x is given by

$P(x_i) = P(x_i)$	0	k $\left(\frac{1}{10}\right)$	$2k$ $2\left(\frac{1}{10}\right)$	$2k$ $2\left(\frac{1}{10}\right)$	$3k$ $3\left(\frac{1}{10}\right)$	k^2 $\left(\frac{1}{10}\right)^2$	$2k^2$ $2\left(\frac{1}{10}\right)^2$	$7k^2+k$ $7\left(\frac{1}{10}\right)^2+\frac{1}{10}$
$F_X(x) = F(x)$	0	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{5}{10}$	$\frac{8}{10}$	$\frac{81}{100}$	$\frac{83}{100}$	1

e) mean $(\mu) = E(X) = \sum P_i x_i$

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

$= 3.66$

f) variance $(\sigma^2) = \sum P_i x_i^2 - \mu^2$

$= 0^2 \times 0 + \left(1^2 \times \frac{1}{10}\right) + \left(2^2 \times \frac{2}{10}\right) + \left(3^2 \times \frac{2}{10}\right) + \left(4^2 \times \frac{3}{10}\right) + \left(5^2 \times \frac{1}{100}\right) + \left(6^2 \times \frac{2}{100}\right) + \left(7^2 \times \frac{17}{100}\right) - (3.66)^2$

$= 16.8 - (3.66)^2$

$= 3.4044$

standard deviation $= \sqrt{\sigma^2} = \sqrt{3.4044} = 1.84$

2. Let x denotes the min. of 2 no's that appear when a pair of dice thrown once. Determine 1) DPDF

2. Expectation 3. variance & standard deviation.

when 2 dice are thrown then the total no. of outcomes are 6^2

$n(S) = 6^2 = 36 =$ (1,1) (1,2) (1,3) (1,4) (1,5) (1,6)
(2,1) (2,2) (2,3) (2,4) (2,5) (2,6)
(3,1) (3,2) (3,3) (3,4) (3,5) (3,6)
(4,1) (4,2) (4,3) (4,4) (4,5) (4,6)
(5,1) (5,2) (5,3) (5,4) (5,5) (5,6)
(6,1) (6,2) (6,3) (6,4) (6,5) (6,6)

Let the RV X assign the minimum no. in x then the sample space S =

1	1	1	1	1	1
1	2	2	2	2	2
1	2	3	3	3	3
1	2	3	4	4	4
1	2	3	4	5	5
1	2	3	4	5	6

The minimum no's are 1, 2, 3, 4, 5, 6

The probability of getting one

The no. of favourable outcomes to get 1 n(1) = (1,1) (1,2) (1,3) (1,4) (1,5) (1,6) (2,1) (3,1) (4,1) (5,1) (6,1) = 11

$P(X=1) = \frac{n(1)}{n(S)} = \frac{11}{36}$

The ~~proba~~ no. of favourable outcomes to get 2

n(2) = (2,2) (2,3) (2,4) (2,5) (2,6) (4,2) (3,2) (5,2) (6,2) = 9

$P(X=2) = \frac{n(2)}{n(S)} = \frac{9}{36}$

The no. of favourable outcomes to get 3

n(3) = (3,3) (3,4) (3,5) (3,6) (4,3) (5,3) (6,3) = 7

$$P(X=3) = \frac{n(3)}{n(S)} = \frac{7}{36}$$

The no. of favourable outcomes to get 4

$n(4) = (4,4)\ (4,5)\ (4,6)\ (5,4)\ (6,4) = 5$

$$P(X=4) = \frac{n(4)}{n(S)} = \frac{5}{36}$$

The no. of favourable outcomes to get 5

$n(5) = (5,5)\ (5,6)\ (6,5) = 3$

$$P(X=5) = \frac{n(5)}{n(S)} = \frac{3}{36}$$

The no. of favourable outcomes to get 6

$n(6) = (6,6) = 1$

$$P(X=6) = \frac{n(6)}{n(S)} = \frac{1}{36}$$

i) Discrete probability distribution function

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

ii) $E(X) = \mu = \Sigma P_i x_i$

$$= \left(1 \times \frac{11}{36}\right) + \left(2 \times \frac{9}{36}\right) + \left(3 \times \frac{7}{36}\right) + \left(4 \times \frac{5}{36}\right) + \left(5 \times \frac{3}{36}\right) + \left(6 \times \frac{1}{36}\right)$$

$= 2.527$

iii) variance $\sigma^2 = \Sigma P_i x_i^2 - \mu^2$

$$= 1^2 \times \left(\frac{11}{36}\right) + 2^2 \times \left(\frac{9}{36}\right) + 3^2 \times \left(\frac{7}{36}\right) + 4^2 \times \left(\frac{5}{36}\right) + 5^2 \times \left(\frac{3}{36}\right) + 6^2 \times \left(\frac{1}{36}\right)$$

$= \cancel{0.44830}\ 8.3611 - (2.527)^2 = 1.97538$

standard deviation $(\sigma) = \sqrt{\sigma^2} = 1.405$

Find the mean and variance by uniform PDF given by $f(x) = \frac{1}{n}$

$x = 1, 2, \ldots\ n$

X	1	2	- - - -	n
f(x)	1/n	1/n	- - - - - -	1/n

mean $(\mu) = E(X) = \Sigma P_i x_i$

$$= 1 \times \frac{1}{n} + 2 \times \frac{1}{n} + \cdots + n \times \frac{1}{n}$$

$$= \frac{1}{n}[1+2+\cdots+n]$$

$$= \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}$$

variance $(\sigma^2) = \sum P_i x_i^2 - \mu^2$

$$= 1^2\left(\frac{1}{n}\right) + 2^2\left(\frac{1}{n}\right) + 3^2\left(\frac{1}{n}\right) + \cdots + n^2\left(\frac{1}{n}\right)$$

$$= \frac{1}{n}\left[1^2+2^2+3^2+\cdots+n^2\right]$$

$$= \frac{1}{n} \cdot \left[\frac{n(n+1)(2n+1)}{6}\right] = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2$$

3. A sample space of 4 items are selected at random from a box containing 12 items of which 5 are defective. Find the expected no. of defective items.

Let x denotes the no. of defective items among 4 items drawn from 12 items.

Total no. of items = 12

No. of defective items = 5

No. of good items = 7.

No. of possible expected defective items = $12C_4$

x	0	1	2	3	4

then $P(x=0) = P$ (No. defective)

$$= \frac{7C_4}{12C_4} = \frac{\frac{7!}{4!(7-4)!}}{\frac{12!}{4!(12-4)!}} = \frac{7}{99}$$

$P(x=1) = P$ (1 defective & 3 good)

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

$P(x=2) = P$ (2 defective & 2 good)

$$= \frac{5C_2 \times 7C_2}{12C_4} = \frac{14}{33}$$

$P(x=3) = P$ (3 defective & 1 good)

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

$P(X=4) = P$ (4 are defective)

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

X	0	1	2	3	4
$F(x_i)$ $P(X \le 4)$	$7/99$	$\frac{35}{99}$	$\frac{42}{99}$	$\frac{14}{99}$	$\frac{1}{99}$

The expected no. of defective items $E(X) = \sum p_i x_i$

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

$= 1.66$

4. Let X denotes the maximum of the 2 no's that appear when a pair of dice are thrown once. Determine the maximum probability i. DPDF ii) Expectation iii. variance & standard deviation.

When 2 dice are thrown then the total no. of outcomes are $n(S) = 6^2 = 36 =$ (1,1) (1,2) (1,3) (1,4) (1,5) (1,6)

(2,1) (2,2) (2,3) (2,4) (2,5) (2,6)

(3,1) (3,2) (3,3) (3,4) (3,5) (3,6)

(4,1) (4,2) (4,3) (4,4) (4,5) (4,6)

(5,1) (5,2) (5,3) (5,4) (5,5) (5,6)

(6,1) (6,2) (6,3) (6,4) (6,5) (6,6)

Let the RV X assign the maximum no. in X then the sample space $S =$

1	2	3	4	5	6
2	2	3	4	5	6
3	3	3	4	5	6
4	4	4	4	5	6
5	5	5	5	5	6
6	6	6	6	6	6

The maximum no's are 1, 2, 3, 4, 5, 6

The no. of favourable outcomes to get 1

$n(1) = (1,1)$

$$P(x=1) = \frac{n(1)}{n(S)} = \frac{1}{36}$$

The no. of favourable outcomes to get 2

$n(2) = (1,2)\ (2,1)\ (2,2) = 3$

$$P(x=2) = \frac{n(2)}{n(S)} = \frac{3}{36}$$

The no. of favourable outcomes to get 3

$n(3) = (1,3)\ (2,3)\ (3,1)\ (3,2)\ (3,3) = 5$

$$P(x=3) = \frac{n(3)}{n(S)} = \frac{5}{36}$$

The no. of favourable outcomes to get 4

$n(4) = (1,4)\ (2,4)\ (3,4)\ (4,1)\ (4,2)\ (4,3)\ (4,4) = 7$

$$P(x=4) = \frac{n(4)}{n(S)} = \frac{7}{36}$$

The no. of favourable outcomes to get 5

$n(5) = (1,5)\ (2,5)\ (3,5)\ (4,5)\ (5,1)\ (5,2)\ (5,3)\ (5,4)\ (5,5) = 9$

$$P(x=5) = \frac{n(5)}{n(S)} = \frac{9}{36}$$

The no. of favourable outcomes to get 6

$n(6) = (1,6)\ (2,6)\ (3,6)\ (4,6)\ (5,6)\ (6,1)\ (6,2)\ (6,3)\ (6,4)\ (6,5)\ (6,6)$
$= 11$

$$P(x=6) = \frac{n(6)}{n(S)} = \frac{11}{36}$$

i. Discrete probability distribution function

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

ii. $E(x) = \mu = \Sigma p_i x_i$

$$= \left(1 \times \frac{1}{36}\right) + \left(2 \times \frac{3}{36}\right) + \left(3 \times \frac{5}{36}\right) + \left(4 \times \frac{7}{36}\right) + \left(5 \times \frac{9}{36}\right) + \left(6 \times \frac{11}{36}\right)$$

$= 4.472$

iii. variance $\sigma^2 = \Sigma p_i x_i^2 - \mu^2$

$$= 1^2\times\left(\frac{1}{36}\right)+\left(2^2\times\frac{3}{36}\right)+\left(3^2\times\frac{5}{36}\right)+\left(4^2\times\frac{7}{36}\right)+\left(5^2\times\frac{9}{36}\right)+\left(6^2\times\frac{11}{36}\right)$$

$$= 21.972-(4.472)^2$$

$$= 1.9734$$

Standard deviation $(\sigma) = \sqrt{\sigma^2} = \sqrt{1.9734} = 1.404$

Probability Density Distribution Function (PDF):

It is defined as the derivative of the PDF of $F_x(x)$

i.e $= f_x(x) \Rightarrow \frac{d}{dx}[F_x(x)] = f_x(x)$

$f(x) \geq 0$

$F_x(x)$ – PMF/DPDF

$f_x(x)$ – PDF/CPDF

* continuous pro PDF (or) probability density function.

for a continuous RV X a function $f(x)$ is said to be probability Density function.

i. $f(x) \geq 0, \forall x \in R$

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

iii. The probability $P(E)$ is given by $P(E) = \int_E f(x)dx$ is called well defined for any event E.

Let $f(x)$ be the P Density F of rv x then

mean of a distribution is given by $\mu = E(X) = \int_{-\infty}^{\infty} x f(x)dx$

→ If X is defined from A to B then $\mu = E(X) = \int_a^b x f(x)dx$

→ Variance $(\sigma^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$

→ standard deviation $(\sqrt{\sigma^2}) = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2$

Note: mean (or) Expectation of any function $\phi(x)$ is given by

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

Median: In case of continuous distribution median is the point which divides the total Area into 2 equal parts thus X is defined from a to b and m is the median then

$$\int_a^M f(x)dx = \int_M^b f(x)dx = \frac{1}{2}$$

Mode:

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

Mean deviation:

mean deviation about the mean μ is given by

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

If X is a continuous Random variable and $Y=ax+b$ then

** i. $E(Y) = aE(x)+b$

** ii. $V(Y) = a^2[V(x)]$ $V \rightarrow$ variance , a & b constants.

1. Probability Density Function of a RV X is $f(x)=\frac{1}{2}\sin x$, $0\le x\le\pi$ $\Rightarrow$ $f(x) = \begin{cases}\frac{1}{2}\sin x & 0\le x\le \pi\\ 0 & \text{otherwise}\end{cases}$

Find the probability of mean, median and mode. Find the probability 0 and $\pi/2$.

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

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

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

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

$$=\frac{1}{2}\int x\sin x\,dx$$

$$=\frac{1}{2}\left[x(-\cos x)-\int 1.(-\cos x)\,dx\right]_0^{\pi}$$

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

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

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

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

$$f'(x)=0 \Rightarrow \frac{1}{2}\cos x=0$$

$$\cos x=0 \Rightarrow \cos x=\cos\pi/2$$

$$x=\pi/2$$

$f''(x) < 0$

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

$= -\frac{1}{2}\sin\left(\frac{\pi}{2}\right) = \frac{-1}{2} < 0$

$\therefore$ mode of the function $f(x)$ is maximum at $x = \pi/2$

median:

Let 'M' is the median of the distribution then

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

$\Rightarrow \int_a^M f(x)\,dx = \frac{1}{2}$

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

$\Rightarrow \frac{1}{2}\int_a^M \sin x\,dx = \frac{1}{2} \quad \Rightarrow [-\cos x]_a^M = 1$

$\Rightarrow -\cos M + \cos a = 1 \quad \because a = 0,\ b = \pi$

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

$-\cos M + 1 = 1 \quad \Rightarrow +\cos M = 0$

$+\cos M = \cos \pi/2$

$M = \pi/2$

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

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

$= \frac{1}{2}\int_0^{\pi/2} \sin x\,dx = \frac{1}{2}[-\cos x]_0^{\pi/2}$

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

2. A continuous RV x has the probability density function

$f(x) = \begin{cases} kxe^{-\lambda x}, & x \ge 0,\ \lambda = 0 \\ 0 & \text{otherwise} \end{cases}$

Determine i. k ii. mean iii. variance

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

~~$\int_0^\infty k x e^{-\lambda x} = 1$~~ $\Rightarrow \int_{-\infty}^{0} f(x)\,dx + \int f(x)\,dx = 1$

$0 + \int_0^\infty kx.e^{-\lambda x}\,dx = 1$

$\Rightarrow k\int_0^\infty x e^{-\lambda x}\,dx = 1$

$\Rightarrow k\left[x.\frac{e^{-\lambda x}}{-\lambda} - \int 1.\frac{e^{-\lambda x}}{-\lambda}\,dx\right]_0^\infty = 1$

$e^{-\infty} = 0$

$e^{\infty} = \infty$

$e^{0} = 1$

$$K\left[-x.\frac{e^{-\lambda x}}{\lambda}-\frac{e^{-\lambda x}}{\lambda^2}\right]_0^\infty = 1$$

$$\Rightarrow K\left[-\infty.\frac{e^{-\lambda\infty}}{\lambda}-\frac{e^{-\lambda\infty}}{\lambda}-0.\frac{e^{0}}{\lambda^2}+\frac{e^0}{\lambda^2}\right]=1$$

$$\Rightarrow K\left[\frac{1}{\lambda^2}\right]=1$$

$$K=\lambda^2$$

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

$$= \int_{\infty}^{0} x f(x)\,dx + \int_0^\infty x f(x)\,dx$$

$$= 0+\int_0^\infty x^2.\lambda^2 x e^{-\lambda x}dx$$

$$= \lambda^2\int_0^\infty x^2.\lambda e^{-\lambda x}dx$$

$$= \lambda^2\left[x^2.\frac{e^{-\lambda x}}{-\lambda}-\int 2x.\frac{e^{-\lambda x}}{-\lambda}dx\right]_0^\infty$$

$$= \lambda^2\left[-\frac{x^2}{\lambda}e^{-\lambda x}+\frac{2}{\lambda}\int x.e^{-\lambda x}dx\right]_0^\infty$$

$$= \lambda^2\left[\frac{-x^2.e^{-\lambda x}}{\lambda}+\frac{2}{\lambda}\left[-x.\frac{e^{-\lambda x}}{\lambda}-\int\frac{e^{-\lambda x}}{-\lambda}dx\right]\right]_0^\infty$$

$$= \lambda^2\left[\frac{-x^2.e^{-\lambda x}}{\lambda}-\frac{2}{\lambda^2}x.e^{-\lambda x}-\frac{2}{\lambda^3}.e^{-\lambda x}\right]_0^\infty$$

$$\lambda^2\left[0+\frac{2}{\lambda^3}\right]=\frac{2\lambda^2}{\lambda^3}=\frac{2}{\lambda}$$

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

$$= \int_{-\infty}^{0} x^2 f(x)\,dx + \int_0^\infty x^2 f(x)\,dx - \mu^2$$

$$= \int_0^\infty x^2\lambda^2 x.e^{-\lambda x}dx$$

$$= \lambda^2\int_0^\infty x^3 e^{-\lambda x}dx$$

$$= \frac{2}{\lambda^2}$$

3. The probability density function $f(x)$ of a continuous RV X is given by $f(x) = c.e^{-|x|}$, $-\infty < x < \infty$

P/T $c = \frac{1}{2}$ and find the mean, variance and probability i.e variate and also lies b/w 0 and 4.

We know that by continuous function

$\int_{-\infty}^{\infty} f(x)dx = 1, \Rightarrow \int_{-\infty}^{\infty} c.e^{-|x|} = 1$

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

0 if f is odd

$\Rightarrow c.\int_{-\infty}^{\infty} e^{-|x|} = 1$

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

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

$\Rightarrow 2c.[-e^{-\infty} + e^{0}] = 1$

$= 2c\ [0 + \frac{1}{\infty}] = 1 \Rightarrow c = \frac{1}{2}$

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

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

$= 0$ $\because f(x)$ is an odd function

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

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

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

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

$= 2.\left[x^2.\frac{e^{-x}}{-1} - \int 2x.\frac{e^{-x}}{-1}dx\right]_{0}^{\infty}$

$= \left[-x^2.e^{-x} + 2.\int x.e^{-x}dx\right]_{0}^{\infty}$

$= \left[-x^2.e^{-x} + 2\left[x.\frac{e^{-x}}{-1} - \int 1.\frac{e^{-x}}{-1}\right]\right]_{0}^{\infty}$

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

$= \left[-\infty^2.e^{-\infty} - 2.\infty.e^{-\infty} - 2.e^{-\infty} + 0^2.e^{0} + 2.0.e^{0} + 2.e^{0}\right]$

$= [0 + 0^2 + 2.0 + 2] = 2$

$P(0 \le x \le 4) = \int_{0}^{4} f(x)\,dx$

$= \int_{0}^{4} c.e^{-|x|}dx$

$= \frac{1}{2}.\int_{0}^{4} e^{-x}dx$

$$= \frac{1}{2}\left[\frac{e^{-x}}{-1}\right]_0^4 = \frac{-1}{2}\left[e^{-4} - e^{0}\right] = \frac{-1}{2}\left[e^{-4} - 1\right]$$

Probability Distributions (or) Theoretical Distributions:

It is the frequency distribution of an certain event in which frequencies are obtained by mathematical computation.

There are 2 types of Theoretical Distributions

1. Discrete Theoretical Distributions
   a. Binomial distribution.
   b. Poisson distribution
   c. Rectangular distribution
   d. -ve binomial distribution
   e. Geometric distribution.

2. Continuous Theoretical Distributions
   a. Normal distribution
   b. Student-T distribution
   c. Chisquare distribution
   d. F-distribution

* Let E be the event and the probability of successive of E is p and failure is q in a single trial and $p+q=1$. This trial is called Bernoulli's trial.

Binomial (or) Bernoulli's Trial: The probability of r successes in n trials in any order is given by

$$P(r) = P(X=r) = nC_r\, p^r\, q^{n-r}, \quad r = 0, 1, 2, \ldots n$$

$$= 0 \quad \text{Otherwise}$$

i.e the Binomial distribution function is given by

$$F_X(x) = P(X \le x) = \sum_{r=0}^{n} nC_r\, p^r . q^{n-r}.$$

The no. of defective bolts containing 9n bolts.

The no. of post graduate in a group of n men.

Conditions:

1. Trials are repeated under identical conditions for a fixed no. of items.
2. There are only 2 possible outcomes.

3. The probability of success in each trial remains constant and does not change trial to trial.

4. The trials are independent.

mean o-

**mean of the binomial distribution $\mu = E(x) = np$

** variance $(x) = \sigma^2 = npq$

mode of the binomial distribution

standard deviation $- \sqrt{npq}$

mode of the binomial distribution: It is the value of x at which P(x) is having max. value.

mode = integral part of $(n+1)p$ if $(n+1)p$ is not an integer $[(n+1)p-1]$

$= [(n+1)p - 1]$, $(n+1)p$ is an integer.

Regerence function:

$$P(r+1) = \frac{(n-r)}{(r+1)} \cdot \frac{p}{q} \cdot P(r)$$

Binomial frequency distribution:

If n independent trials consecute 1 experiment and this is experiment is repeated N times then the frequency of r successes is $N \cdot {}^nC_r \, p^r q^{n-r}$ and the possible no. of successes and their frequencies is called a Binomial frequency distribution and calculated by $N \cdot (q+p)^n$

1. A fair coin is tossed 6 times. Find the probability of getting 4 heads

Given that

A coin is tossed 6 times.

let Assume that p is the probability of getting a head (or) q is the probability of getting a tail

$$P(H) = \frac{1}{2}$$

$$q(T) = \frac{1}{2}$$

$$n = 6$$

Let us Assume that no. of success r i.e no. of heads 4

$P(X = r) = P(X=4) = nC_r \cdot p^r \cdot q^{n-r}$

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

$= \frac{15}{64} = 0.234$

2. 10 coins are thrown simultaneous. Find the probability of getting a) atleast 6 heads b) atleast 8 heads.

No. of coins $(n) = 10$

Let p is the probability of getting head i.e $p = P(H) = \frac{1}{2}$

q is the probability of getting tail $\Rightarrow q = P(T) = \frac{1}{2}$

$P(x=r) = P(x=6) = nC_r \cdot p^r \cdot q^{n-r}$

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

$= \frac{105}{512} = 0.205$

$P(X \geq r) = P(r \geq 6)$

$= P(x=6) + P(x=7) + P(x=8) + P(x=9) + P(x=10)$

$= 10C_6 \left(\frac{1}{2}\right)^6 \cdot \left(\frac{1}{2}\right)^{10-6} + 10C_7 \left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^{10-7} + 10C_8 \left(\frac{1}{2}\right)^8 \left(\frac{1}{2}\right)^{10-8} +$

$10C_9 \left(\frac{1}{2}\right)^9 \cdot \left(\frac{1}{2}\right)^{10-9} + 10C_{10} \left(\frac{1}{2}\right)^{10} \left(\frac{1}{2}\right)^{10-10}.$

$= \frac{105}{512} + \frac{15}{128} + \frac{45}{1024} + \frac{5}{512} + \frac{1}{1024}$

$= \frac{193}{512} = 0.376$

$P(x \geq r) = P(r \geq 8)$

$= P(x=8) + P(x=9) + P(x=10)$

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

$= \frac{45}{1024} + \frac{5}{512} + \frac{1}{1024}$

$= \frac{7}{128} = 0.054$

3. Determine the binomial distribution for which mean μ &

variant is 3. Also find mode.

mean B.D $(\mu) = np = 4$ — ①

variance B.D$(\sigma^2) = npq = 3$ — ②

$$\frac{②}{①} = \frac{npq}{np} = \frac{3}{4}$$

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

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

$$np = 4$$

$$n\left(\frac{1}{4}\right) = 4 \Rightarrow n = 16$$

$$P(x=r) = nC_r \cdot p^r \cdot q^{n-r}$$

$$= 16C_r \cdot \left(\frac{1}{4}\right)^r \cdot \left(\frac{3}{4}\right)^{16-r}$$

mode:

$$(n+1)p \Rightarrow (16+1)\left(\frac{1}{4}\right)$$

$$= 4.25$$

$$[(n+1)p-1] = \left[(16+1)\left(\frac{1}{4}\right)-1\right] = 4.25-1 = 3.25$$

mode of the B.D = 4

4. The mean & variance of a distribution is 4 & $\frac{4}{3}$ then find $P(x \geq 1)$.

Given that,

mean B.D$(\mu) = np = 4$ — ①

variance $(\sigma^2) = npq = \frac{4}{3}$ — ②

$$\frac{②}{①} = \frac{npq}{np} = \frac{4/3}{4}$$

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

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

$$np = 4$$

$$n\left(\frac{2}{3}\right) = 4 \Rightarrow n = 6$$

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

$p(x=r) = nC_r \cdot p^r q^{n-r}$

$= 6C_1\left(\frac{2}{3}\right)^1\left(\frac{1}{3}\right)^{6-1} + 6C_2\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^{6-2} + 6C_3\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^{6-3} +$

$6C_4\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^{6-4} + 6C_5\left(\frac{2}{3}\right)^5\left(\frac{1}{3}\right)^{6-5} + 6C_6\left(\frac{2}{3}\right)^6\left(\frac{1}{3}\right)^{6-6}$

$= \frac{4}{243} + \frac{20}{243} + \frac{160}{729} + \frac{80}{243} + \frac{64}{243} + \frac{64}{729}$

$= 0.998$

5. Out of 800 families with 5 children each how many would you expect to have a) 3 boys b) 5 girls c) either 2 or 3 boys d) atleast 1 boy.

Assume equal probability for boys & girls.

Let P = probability of having a boy in each family

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

$q = P(G) = 1/2$

probability of getting children $n=5$

$P(x=r) = nC_r \cdot p^r \cdot q^{n-r}$

$= 5C_r \cdot \left(\frac{1}{2}\right)^r \left(\frac{1}{2}\right)^{5-r}$

$= 5C_r \left(\frac{1}{2}\right)^5$

a) Let r = probability of having no. of boys

Then the probability of getting 3 boys

$r=3$

$P(x=r) = P(x=3) = 5C_3\left(\frac{1}{2}\right)^5$

$= \frac{5}{16} = 0.3125$

Thus for 800 families the probability of no. of families having 3 boys

$= 800 \times 0.3125 = 250$

b) 5 girls probability of having 5 girls (or) no boys

$P(x=r) = P(x=5) = 5C_5 \cdot \left(\frac{1}{2}\right)^5$

$= \frac{1}{32} = 0.03125$

Thus the probability of having 5 girls (or) no boys

$= 0.031 \times 800 = 25$ families

c) probability of having 2 (or) 3 boys

$P(x = 2 \text{ or } 3) = P(x=2) + P(x=3)$

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

$= 0.625$

d) probability of getting atleast 1 boy

$P(x=1) = 0.625 \times 800 = 500$ families

Let r is = no. of boys

$P(x=r) = P(r \geq 1)$

$= P(x=1) + P(x=2) + P(x=3) + P(x=4) + P(x=5)$

$= 5C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^{5-1} + 5C_2 \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^{5-2} + 5C_3 \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^{5-3} +$

$5C_4 \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^{5-4} + 5C_5 \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right)^0$

$= 0.156 + 0.312 + 0.312 + 0.156 + 0.03$

$= 0.966$

$= 0.96875 \times 800 = 775$

6. 20% of items are produced from a factory are defective find the probability that in a sample 5 is chosen at random i. None is defective ii. one is defective iii. $P(1 < x < 4)$

7. Fit a B.D for the values

x	0	1	2	3	4	5	6
f	2	14	20	34	22	8	

consider from the given data $n=5$

$$\Sigma f = N = 100$$

$$\text{mean} = \frac{\Sigma f_i x_i}{\Sigma f = N}$$

$$= \frac{(0\times2)+(1\times14)+(2\times20)+(3\times34)+(4\times22)+(5\times8)}{100}$$

$$= \frac{14+40+102+88+40}{100} = 2.84$$

$$\text{mean } (\mu) = np = 2.84$$

$$5P = 2.84$$

$$P = 0.568$$

$$q = 1-P = 1-0.568 = 0.432$$

Hence Binomial frequency Distribution is $N(q+P)^n$

$= 100\,[0.432+0.568]^5$

$(q+p)^n = nC_0 q^n p^0 + nC_1 q^{n-1} p^1 + nC_2 q^{n-2} p^2 + \cdots$

$= 100\,[5C_0 (0.432)^5 (0.568)^0 + 5C_1 (0.432)^4 (0.568)^1 +$
$5C_2 (0.432)^3 (0.568)^2 + 5C_3 (0.432)^2 (0.568)^3 + 5C_4 (0.432)^1 (0.568)^4 +$
$5C_5 (0.432)^0 (0.568)^5]$

$= 100\,[0.015 + 0.098 + 0.260 + 0.341 + 0.224 + 0.059]$

$= 1.5 + 9.89 + 26.01 + 34.19 + 22.48 + 5.90$

$= 100$

At last Hence the B.D Table is

x	0	1	2	3	4	5
f	2	14	20	34	22	8
Expected (or) Theoretical frequencies	2	10	26	34	22	6

** 8. Four coins are tossed 160 times, the no. of times 'x is head occur' given below. Fit a B.D to this data and assume that the hypothesis coins are unbiased.

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

Given that,

The coins are unbiased

probability of getting head $= \frac{1}{2}$

probability of getting Tail $= \frac{1}{2}$

$N = 160 = f = $ No. of times coin tossed.

$n = 4$

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

$P(x=0) = 4C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^{4-0} = 0.0625$

$P(x=1) = 4C_1 (0.5)^1 (0.5)^3 = 0.25$

$P(x=2) = 4C_2 (0.5)^2 (0.5)^2 = 0.375$

$P(x=3) = 4C_3 (0.5)^3 (0.5)^1 = 0.25$

$P(x=4) = 4C_4 (0.5)^4 (0.5)^0 = 0.0625$

No. of heads x	No. of Times f	Probability $P(x)$	Expected (or) Theoretical frequency $f(x) = N P(x)$
0	8	0.0625	10
1	34	0.25	40
2	69	0.375	60
3	43	0.25	40
4	6	0.0625	10
	160		160

9. The probability that a man hitting a target is $\frac{1}{3}$. If he fires 6 times. Find The probability that he fires i) At the most 5 times ii) Exactly ones. iii) Atleast two times.

Given that,

The probability of hitting a target $p = \frac{1}{3}$

No. of Trials $= n = 6$

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

i) probability of firing at most 5 times

$$P(x \le x) = P(x \le 5) = 1 - P(x=6) = 1 - nC_x \, p^x q^{n-x}$$

$$= 1 - \left[6C_6 \left(\frac{1}{3}\right)^6 \left(\frac{2}{3}\right)^{6-6}\right]$$

$$= 0.9986$$

ii) $P(x=1) = 6C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^{6-1} = 0.263$

iii) $P(x \ge 2) = P(x=2) + P(x=3) + P(x=4) + P(x=5) + P(x=6)$ (or)

$$= 1 - P(x<2)$$

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

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

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

$$= 1 - 0.3511 = 0.6489$$

10. A coin is biased in a way that head is twice "likely to occur" as a tail. If the coin is tossed 3 times find the probability of getting two tails and one head.

Given that,

No. of Trials $n = 3$

$P(H) + P(T) = 1$

$2P(T) + P(T) = 1$

$3P(T) = 1$

$P(T) = \frac{1}{3}$

$P(H) = q = 1 - \frac{1}{3} = \frac{2}{3}$

$P(x=x) = P(x=2) = 3C_2 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^1 = 0.222$

Poisson Distribution:

A RV 'x' is said to be follow poisson distribution. If it Assumes only non -ve values and its probability density function is given by $P(x,\lambda) = P(x=x) = \frac{e^{-\lambda} \lambda^x}{x!}, x = 0,1,2,\ldots\infty$

0 otherwise

Here $\lambda > 0$ is a parameter of the distribution and the poisson distribution function is $F_x(x) = P(x \le x) = P(x=r) = e^{-\lambda} \sum_{r=0}^{x} \frac{\lambda}{r!}$, $\lambda > 0$

Derivation of poisson distribution:

It can be derived as a limiting case of the B.D under the conditions that

i. No. of Trials (n is very large)

ii. The probability of success p (is very small (very close to 0).

iii. $np = \lambda$ is a finite value.

Mean of the poisson distribution:

** mean $(\mu) = \lambda$

** variance $(\sigma^2) = \lambda$

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

$\mu = \sigma^2 = x$

→ mode of the poisson Distribution lies b/w $\lambda - 1$ & λ

Recurrence relation of the poisson distribution

$$P(x+1) = \frac{\lambda}{x+1} P(x) \quad (or)$$

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

1. A Hospital switchboard receives an average of 4 emergency calls in 10 minutes interval. what is the probability that i) There are at most 2 emergency calls in a 10 minutes interval. ii) There are exactly 3 emergency " " "

Given that

λ = Avg. no. of calls = 4 calls in a 10 minutes

$$P(X=x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

i) probability of ~~atleast~~ most 2 emergency calls

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

$$= \frac{e^{-4} \cdot 4^0}{0!} + \frac{e^{-4} \cdot 4^1}{1!} + \frac{e^{-4} \cdot 4^2}{2!}$$

$$= e^{-4}(0.238)$$

ii) probability of exactly 3 emergency calls.

$$P(X=x=3) = \frac{e^{-4} \cdot 4^3}{3!} = 0.195$$

2. A manufacture of cotter pins knows that 5% product is defective. pins. (are sold in boxes of 100. the gene guarantee is that not more than 10 pins will be defective. what is the approximate probability that a box will fail to meet the guarantee quality.

Let p is the probability no. of defective pins

$$P = 5\% = \frac{5}{100} = 0.05$$

$$q = 1-p = 1-0.05 = 0.95$$

no. of pins $n = 100$

Let mean of the poisson distribution $= np = \lambda$

$100(0.05) = \lambda$

$\lambda = 5$

$$P(X=x) = \frac{e^{-\lambda} \cdot \lambda^{x}}{x!}$$

P(a box will fail to meet the guarantee quality) $= 1 - P(X \leq 10)$

$= 1 - [P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5) + P(X=6) + P(X=7) + P(X=8) + P(X=9) + P(X=10)]$

$$= 1 - \left[\frac{e^{-5} \cdot (5)^0}{0!} + \frac{e^{-5} \cdot 5^1}{1!} + \frac{e^{-5} \cdot 5^2}{2!} + \frac{e^{-5} \cdot 5^3}{3!} + \frac{e^{-5} \cdot 5^4}{4!} + \frac{e^{-5} \cdot 5^5}{5!} + \frac{e^{-5} \cdot 5^6}{6!} + \frac{e^{-5} \cdot 5^7}{7!} + \frac{e^{-5} \cdot 5^8}{8!} + \frac{e^{-5} \cdot 5^9}{9!} + \frac{e^{-5} \cdot 5^{10}}{10!}\right]$$

$= 0.0137$

3. If X is a poisson variate such that $3P(X=4) = \frac{1}{2} P(X=2) + P(X=0)$. Find i) mean of $x(\lambda)$ ii) $P(X \leq 2)$

i) Given that

i) $3P(X=4) = \frac{1}{2} P(X=2) + P(X=0)$

$$\Rightarrow 3 \frac{e^{-\lambda} \cdot \lambda^4}{4!} = \frac{1}{2} \cdot \frac{e^{-\lambda} \cdot \lambda^2}{2!} + \frac{e^{-\lambda} \cdot \lambda^0}{0!}$$

$$\Rightarrow e^{-\lambda}\left[\frac{3\lambda^4}{24}\right] = e^{-\lambda} \cdot \left[\frac{\lambda^2}{4} + 1\right]$$

$$\Rightarrow \frac{3\lambda^4}{24} = \frac{\lambda^2}{4} + 1 \quad \Rightarrow \frac{3\lambda^4}{24} = \frac{\lambda^2 + 1}{4}$$

$\{ \Rightarrow 3\lambda^4 = 6(\lambda^2 + 1) \Rightarrow \cancel{3}\lambda^4 = 2(\lambda^2 + 1)$

$\Rightarrow \lambda^4 - 2\lambda^2 + 1 \Rightarrow \lambda^4 - 2\lambda^2 - 2 = 0 \}$

$$\frac{\lambda^4}{8} = \frac{\lambda^2}{4} + 1 \Rightarrow \lambda^4 - 2\lambda^2 - 8 = 0$$

By synthetic division

2	1	0	-2	0	-8
	0	2	4	4	8
-2	1	2	2	4	0
	0	-2	0	-4	
	1	0	2	0	

$\lambda^2 + 2 = 0 \quad \Rightarrow \lambda^2 = -2$

$\lambda^2 = 2i^2$

$\lambda = \pm i\sqrt{2}$

$\lambda = -2, 2, \pm i\sqrt{2} \quad (\because \lambda > 0)$

$\therefore \lambda = 2$

ii) $P(x \le 2) = P(x=0) + P(x=1) + P(x=2)$

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

$= e^{-2} \cdot [1 + 2 + 2] = e^{-2} \cdot 5 = 0.676$

4. Fit a poisson Distribution for the following data and calculate the expected frequency.

x	0	1	2	3	4
$f(x)$	109	65	22	3	1

$N = \Sigma f_i = 200$

$\text{mean } (\mu) = \lambda = \frac{\Sigma f_i x_i}{N}$

$= \frac{(109 \times 0) + (1 \times 65) + (2 \times 22) + (3 \times 3) + (4 \times 1)}{200}$

$= \frac{65 + 44 + 9 + 4}{200} = \frac{122}{200} = 0.61$

$P(x = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$

$P(x=0) = \frac{e^{-0.61} (0.61)^0}{0!} = 0.5433$

$P(x=1) = \frac{e^{-0.61} (0.61)^1}{1!} = 0.3314$

$P(x=2) = \frac{e^{-0.61} (0.61)^2}{2!} = 0.1010$

$P(x=3) = \frac{e^{-0.61} (0.61)^3}{3!} = 0.0205$

$P(x=4) = \frac{e^{-0.61} (0.61)^4}{4!} = 0.00313$

x	Observed frequency $f(x)$	$P(x)$	Expected frequencies $N P(x)$
0	109	0.543	$108.6 \approx 109$
1	65	0.331	$66.2 \approx 66$
2	22	0.101	$20.2 \approx 20$
3	3	0.020	$4 = 4$
4	1	0.003	$0.6 \approx 1$
	200		200

5. Using Recurrence formula. Find the probabilities when $x=0,1,2,3,4,5$. If the mean of the poisson distribution is 3.

Given that,

mean of the poisson distribution $= \lambda = 3$

By the P.D $P(X=x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}, \lambda > 0, x = 0,1, \dots \infty$

$$= \frac{e^{-3} \cdot 3^x}{x!}$$

By recurrence formula we know that

$$P(x+1) = \frac{\lambda}{x+1} P(x)$$

$$= \frac{3}{x+1} P(x)$$

put $x=0 \Rightarrow P(1) = \frac{3}{1} P(0)$

$$P(x=0) = \frac{e^{-3} \cdot 3^0}{0!} = 0.0497$$

$$\Rightarrow P(0+1) = \frac{3}{0+1} P(0)$$

$$= 3(0.049) = 0.149$$

put $x=1 \Rightarrow P(1+1) = \frac{3}{1+1} \cdot P(1)$

$$= \frac{3}{2} \cdot \frac{e^{-3} \cdot 3^1}{1!} = 0.2240$$

put $x=2 \Rightarrow P(2+1) = \frac{3}{2+1} P(2)$

$$= \frac{3}{3} \frac{e^{-3} \cdot 3^2}{2!} = 0.2240$$

put $x=3 \Rightarrow P(3+1) = \frac{3}{3+1} P(3)$

$= \frac{3}{4} \cdot \frac{e^{-3} 3^3}{3!} = 0.1680$

put $x=4 \Rightarrow P(4+1) = \frac{3}{4+1} P(4)$

$= \frac{3}{5} \cdot \frac{e^{-3} 3^4}{4!} = 0.1008$

put $x=5 \Rightarrow P(5+1) = \frac{3}{5+1} P(5)$

$= \frac{3}{6} \cdot \frac{e^{-3} 3^5}{5!} = 0.0504$

put $x=6 \Rightarrow P(6+1) = \frac{3}{6+1} P(6)$

$= \frac{3}{7} \cdot \frac{e^{-3} 3^6}{6!} = 0.0216$

x	$P(x)$	Recurrence $P(x+1) = \frac{\lambda}{x+1} P(x)$
0	0.149	0.224
1	0.224	0.224
2	0.224	0.168
3	0.168	0.100
4	0.100	0.050
5	0.050	0.021

Normal Distribution:

A RV x is said to have normal distribution, if its density function (or) probability distribution function

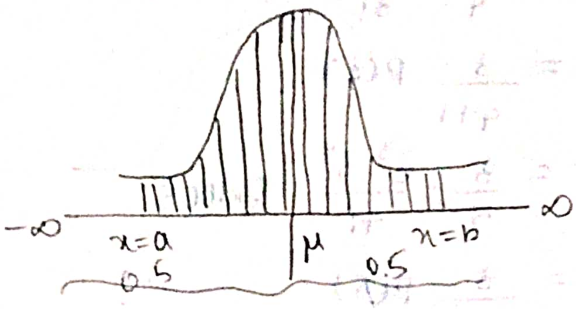
$$f(x,\mu,\sigma) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ (or) } \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} \\ -\infty < x < \infty \\ -\infty < \mu < \infty \end{cases}$$

where μ = mean
σ = standard deviation are the 2 parameters of

Normal distribution

The RV x is said to be a normal rv (or) Normal variate the curve representing the normal distribution is called the

Normal curve and the total area bounded by the curve on the x-axis is '1' i.e $\int_{-\infty}^{\infty} f(x)\,dx = 1$



Mean of the Normal Distribution:

Consider the N.D with (a,σ) are the parameters then

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

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

$$= a$$

Variance of Normal Distribution:

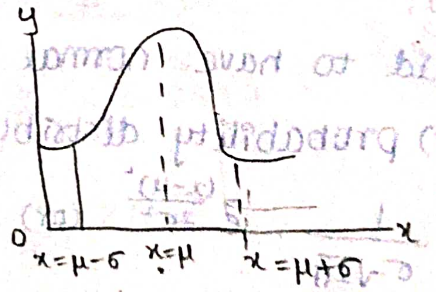
$$\text{Variance } (\sigma^2) = E(x^2) - [E(x)]^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx$$

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

$$= \sigma^2$$

Characteristics of a Normal Distribution:

The graph of the N.D $y=f(x)$ in the x-y plane is known as the normal curve.



The curve is bell shaped and symmetric about the line $x=\mu$ mean, median and mode of the distribution is coincide so normal curve is unimodel.

x-axis is asymptotic curve area of the normal curve lies b/w $\mu-\sigma$, $\mu+\sigma$

If $\mu=0$ & $\sigma=1$ then the distribution is called standard normal distribution. By taking $z=\mu-\sigma$, the standard

Normal curve is form.

The total area under the curve is 1. and the area under this curve is divided into 2 parts

procedure to find probability density of Normal curve:

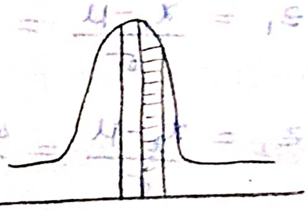
The probability that Normal variate x with mean μ and standard deviation as 'σ' that lies b/w two specific values x_1 & x_2

1. Take $z = \frac{x-\mu}{\sigma}$ and find z_1 & z_2 corresponding to the values x_1 & x_2 respectively

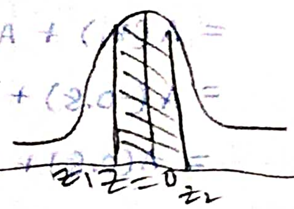
2. probability $(x_1 \le x \le x_2) = P(z_1 \le z \le z_2)$

i. If both z_1 & z_2 are +ve (or) -ve i.e area under the normal curve 0 to z_2 - 0 to z_1

$A(z_2) - A(z_1)$



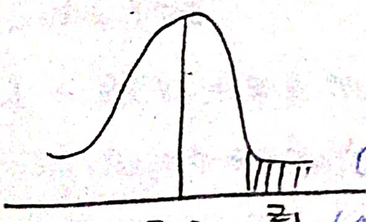
ii. If $z_1 < 0$ & $z_2 > 0$



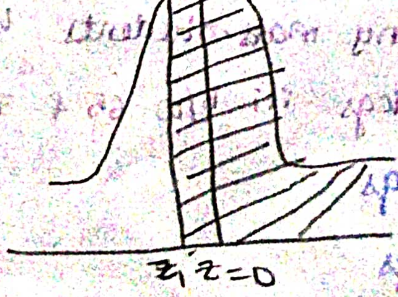
Then $P(x_1 \le x \le x_2) = P(z_1 \le z \le z_2)$

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

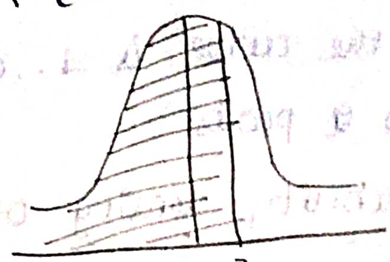
iii. If $z_1 > 0$ then $P(z > z_1) = 0.5 - A(z_1)$



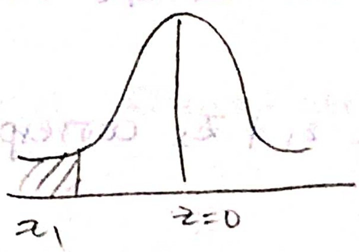
iv. If $z_1 < 0$ then $P(z > z_1) = 0.5 + A(z_1)$



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



vi. If $z_1 \le 0$ then $P(z < z_1) = 0.5 - A(z_1)$



1. If x is a Normal variate with mean 30 & $\sigma = 5$. Find i. $26 \le x \le 40$ ii. $x \ge 45$.

i) Let $x_1 = 26 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma} = \frac{26-30}{5} = -0.8$

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

from case-ii) $P(x_1 < x \le x_2) = P(z_1 \le z \le z_2)$

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

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

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

$= 0.2881 + 0.4772$

$= 0.7656$

$A(0.8) = 0.2881$

$A(2) = 0.4772$

ii) $x = 45 \Rightarrow z_1 = \frac{x-\mu}{\sigma} = \frac{45-30}{5} = 3$

$z_1 = 3 > 0$

$P(x \ge 45) = P(z > z_1)$

$= 0.5 - A(z_1)$

$= 0.5 - A(3.0)$

$= 0.5 - 0.4987 = 0.0013$

2. If the masses of 300 students are normally distributed with mean 68^{kg} & $\sigma = 3$ kg. i) How many students have masses greater than 72 kg ii. ≤ 64 kgs iii. B/w 65 & 71 kgs (inclusive i.e $65 \le 71$)

Given that $\mu = 68$ kgs

$\sigma = 3$ kgs

i) Let $x = 72 \Rightarrow z = \frac{x-\mu}{\sigma} = \frac{72-68}{3} = 1.33$

$\therefore z_1 = 1.33$

$P(x > 72) = P(z > z_1)$

$= 0.5 - A(z_1)$

$= 0.5 - A(1.33)$

$= 0.5 - 0.4082 = 0.0918$

No. of students more than 72 kgs masses is 300×0.0918

$= 27.54 \simeq 28$

ii) $x = 64 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma} = \frac{64-68}{3} = -1.33$

$z_1 = -1.33 < 0$

$P(x \leq 64) = P(z \leq z_1)$

$= 0.5 - A(z_1)$

$= 0.5 - A(-1.33) = 0.5 - A(1.33)$

$= 0.5 - 0.4028 = 0.0918$

No. of students has masses less than 64 kgs $= 300 \times 0.0918$

$= 27.54 \approx 28$

iii) $P(65 \leq x \leq 71)$

$x_1 = 65 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma} = \frac{65-68}{3} = -1$

$z_1 = -1 < 0$

$x_2 = 71 \Rightarrow z_2 = \frac{71-68}{3} = \frac{3}{3} = 1$

$z = 1 > 0$

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

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

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

$= 0.3413 + 0.3413 = 0.6826$

No. of students of masses b/w 65 and 71 kgs $= 300 \times 0.6826$

$= 204.78 \simeq 205$

3. In a N.D 7% of items are under 35 and 89% are under 63. Determine the i) mean and variance of the distribution.

Let μ is the mean of the distribution &

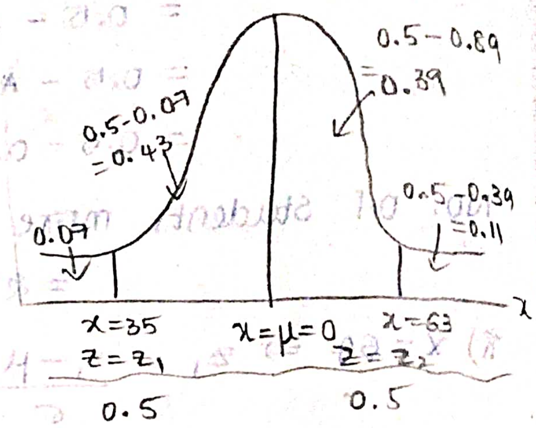
σ is the standard deviation of the distribution

Given that, 7% of items under 35

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

and 89% of items under 63

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



Let $X=35 \Rightarrow z_1 = \frac{x-\mu}{\sigma} = \frac{35-\mu}{\sigma}$ — ①

let $X=63 \Rightarrow z_2 = \frac{x-\mu}{\sigma} = \frac{63-\mu}{\sigma}$ — ②

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

$z_1 = 1.48$ (from the table)

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

$z_2 = 1.23$ (from the table)

① $\Rightarrow -1.48 = \frac{35-\mu}{\sigma}$ — ③

② $\Rightarrow 1.23 = \frac{63-\mu}{\sigma}$ — ④

$-1.48\sigma = 35-\mu$

$1.23\sigma = 63-\mu$

$-2.71\sigma = -28$

$\sigma = 10.33$

④ $\Rightarrow 1.23(10.33) = 63-\mu$

$\mu = 50.29$

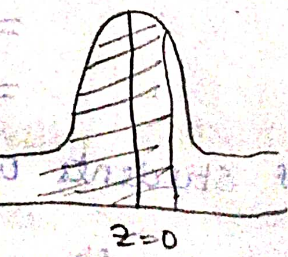
4. The marks obtained in mathematics by 1000 students is N.D with $\mu = 78\%$ & $\sigma = 11\%$. Determine i) ~~Abo~~ How many students got above 90% ii. What was the highest mark obtained by the lowest 10% of the students.

Given that $\mu = 78\% = 0.78$

$\sigma = 11\% = 0.11$

i) Let $x = 90\% = 0.90$

$z_1 = \frac{0.90-0.78}{0.11} = 1.09$



$P(X>90\%) = P(z>z_1)$

from case (8) $= 0.5 - A(z_1)$

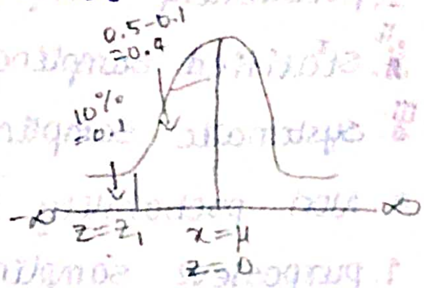
$= 0.5 - A(1.09) = 0.5 - 0.3621$

$= 0.1379$

Hence the no. of students with marks more than 90% is

$= 137.9 \simeq 138$

ii) 0.1 area to the left of z corresponding to the lowest 10% of the students.



$Z = A(z_1) = 0.5 - A(z = z_1, \text{ to } \infty)$

$= 0.5 - 0.1 = 0.4$

$z_1 = -1.29$ (from the table)

$\Rightarrow z_1 = \frac{x - \mu}{\sigma}$

$-1.29(0.11) = x - 0.78$

$x = 0.6381$

∴ The total no. of students with marks obtained by the 10% of the students $= 1000 \times 0.638$

$= 68 = 638 = 63.8\%$

Sampling Distribution:

population: It is the aggregate (or) totality of statistical data forming a subject of investigation.

Ex: The population in India.

Sample (or) sampling: It is the subset of the population and the no. of objects in sample is called size of the population

If the sample size $n < 30$ then it is said to be small sample.

If $n \geq 30$ then it is said to be large small.

Total no. of samples we can select with size of sample n from N.

∴ No. of population without replacement is NC_n

No. of population with replacement is N^n

Sampling Error: In estimating the characteristics of the population instead of enumerating entire population only the individual in the sample are examined.

The error involved in such approximations is known as sampling error.

Different methods of sampling method

1. probability sampling method : i. Random sampling

ii. statified sampling method

iii. systematic sampling

2. Non probability sampling method.

i. purposieve sampling (or) judgement sampling

ii. sequential sampling.

Statistic: It is the real valued function of the random sample.

Statistic is a function of one (or) more ~~r.v~~ r.v not involving any unknown parameter

Parameter: It is a statistical measure based on all the observation of the population.

Ex: μ - mean, σ^2 - variance

Sampling ~~var~~ fluctuation: statistics varies from one sample to another sample, this variation in the value of statistics is called sampling fluctuation.

sample mean: If $x_1, x_2, \ldots x_n$ represent a random sample of size n then the sample mean is defined by $\bar{x} = \frac{\Sigma x_i}{n}$, $\mu = \frac{\Sigma x_i}{N}$.

Sample variance: $\sigma^2 = \frac{(x_i - \mu)^2}{N}$ (population variance)

sample variance $s^2 = \frac{\Sigma (x_i - \bar{x})^2}{n-1}$

Note: mean of the population = mean of the sample

Central ~~line~~ limiting theorem:

If $\bar{x}$ is the mean of the random sample from n taken from a population having the mean (μ) and standard deviation (σ) then $z = \frac{\bar{x} - \mu}{(\sigma/\sqrt{n})}$ is a ~~r.v~~ r.v that approaches to N.D with $n \to \infty$

Sample Error:

It is the S.D of the sample distribution and is denoted by $E = \frac{\sigma}{\sqrt{n}}$

For a finite population N, the standard error of sample mean is. $S.E = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

Standard Error for sample proportion:

$$S.E = \sqrt{\frac{PQ}{n}}\sqrt{\frac{N-n}{N-1}}$$

p represents occuring of sample

where $\frac{N-n}{N-1}$ is called correlation factor.

For a infinite population, the standard error of sample mean $S.E = \frac{\sigma}{\sqrt{n}}$

$$S.E = \sqrt{\frac{PQ}{n}} \quad \text{(sample proportion)}$$

standard Error of sample is $S.E = \frac{\sigma}{\sqrt{2n}}$

standard Error of difference of 2 samples is $\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$

Standard Error of 2 different proportions $P_1-P_2 = \sqrt{\frac{P_1Q_1}{n_1}+\frac{P_2Q_2}{n_2}}$

S.E of 2 sample sizes $S_1-S_2 = \sqrt{\frac{\sigma_1^2}{2n_1}+\frac{\sigma_2^2}{2n_2}}$

1. Find the corelation factor for the following population. is 200, n=5

Given that N = 200, n=5

Corelation factor is $\frac{N-n}{N-1}$

$$= \frac{200-5}{200-1} = 0.9798$$

2. A population consists of 5 no's 2,3,6,8,11. consider all possible samples of sizes 2 which can be drawn with replacement from the population. Find i) mean of population. ii. S.D of population iii. mean of the sampling distribution of means. iv. S.D of sampling distribution of means. i.e standard errors of the mean.

Given that

population size $N = 5$

sample size $n = 2$

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

ii) $\sigma^2 = \frac{\Sigma (x_i - \mu)^2}{N} = \frac{(30-6)^2}{5} =$

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

$= \frac{16+9+4+25}{5} = \frac{54}{5} = 10.8$

$\sigma = \sqrt{10.8} = 3.28$

iii. Total no. of samples with replacement $= N^n$

$= 5^2 = 25$

∴ The samples are (2,2) (2,3) (2,6) (2,8) (2,11)

(3,2) (3,3) (3,6) (3,8) (3,11)

(6,2) (6,3) (6,6) (6,8) (6,11)

(8,2) (8,3) (8,6) (8,8) (8,11)

(11,2) (11,3) (11,6) (11,8) (11,11)

iv. The means of samples are $= \frac{2+2}{2} \quad \frac{2+3}{2} \quad \frac{2+6}{2} \quad \frac{2+8}{2} \quad \frac{2+11}{2}$

$\frac{3+2}{2} \quad \frac{3+3}{2} \quad \frac{3+6}{2} \quad \frac{3+8}{2} \quad \frac{3+11}{2}$

$\frac{6+2}{2} \quad \frac{6+9}{2} \quad \frac{6+6}{2} \quad \frac{6+8}{2} \quad \frac{6+11}{2}$

$\frac{8+2}{2} \quad \frac{8+3}{2} \quad \frac{8+6}{2} \quad \frac{8+8}{2} \quad \frac{8+11}{2}$

$\frac{11+2}{2} \quad \frac{11+3}{2} \quad \frac{11+6}{2} \quad \frac{11+8}{2} \quad \frac{11+11}{2}$

= 2 2.5 4 5 6.5

2.5 3 4.5 5.5 7

4 7.5 6 7 8.5

5 5.5 7 8 9.5

6.5 7 8.5 9.5 11

∴ The mean of the sample distribution of means is

$$\mu_{\bar{x}} = \frac{\text{sum of sample means}}{\text{Total no. of sample means}}$$

$$= \frac{2+2.5+4+5+6.5+2.5+3+4.5+5.5+7+4+7.5+6+7+8.5+5+5.5+7+8+9.5+6.5+7+8.5+9.5+11}{25}$$

$$= \frac{153}{25} = 6.12$$

The S.D of the sample distribution ~~is the vari~~.

The variance of sample distribution

$$\sigma_{\bar{x}}^2 = \frac{\Sigma(x_i-\mu)^2}{N}$$

$$= \frac{(2-6)^2+(2.5-6)^2+(4-6)^2+(6.5-6)^2+(2.5-6)^2+(3-6)^2+(4.5-6)^2+(5.5-6)^2+(7-6)^2+(4-6)^2+(7.5-6)^2+(6-6)^2+(7-6)^2+(8.5-6)^2+(5-6)^2+(5.5-6)^2+(7-6)^2+(8-6)^2+(9.5-6)^2+(6.5-6)^2+(7-6)^2+(8.5-6)^2+(9.5-6)^2+(11-6)^2}{25}$$

$$= \frac{16+12.25+4+0.25+12.25+9+2.25+0.25+1+4+2.25+1+4+\cancel{2.25}+\cancel{1}+6.25+1+0.25+1+4+12.25+0.25+1+6.25+12.25+25}{25}$$

$$= \frac{141.25}{25} = \cancel{5.65} = \frac{134}{25} = 5.36$$

SAMPLING DISTRIBUTION :-

POPULATION :-

It is aggregate of statistical data forming a subject of investigation.

Eg:- 1. population of marks,

2. population of railway stations

3. ~~Size of the~~ po

SIZE OF THE POPULATION :-

The no. of observation in the population is defined to be size of the population.

SAMPLE :-

A sample is a subset of population and the no. of objects in this sample is called size of the sample

LARGE SAMPLE :-

~~Larr~~ If the size of a sample $n \geq 30$, then the sample is called large sample

Small sample :-

If the size of the sample $n < 30$, then the sample is called small sample

NOTE :-

1. The sampling from finite population with replacement can be considered as sampling from infinite population.

2. The sampling from finite population without replacement can be considered as sampling from finite population

STATISTIC :-

It is the real-valued function of the Random variable (R.V).

Eg:- The value of mean and the data in population

The data in sample PTC

PARAMETER :-

A statistical measure based on population is called parameters

Eg:- The statistical constants of population are

1. Mean (μ)

2. Variance (σ^2)

3. Standard deviation (σ)

NOTE :- Statistical constants of the sample are

1. Mean ($\bar{x}$) 2. Variance (s^2) 3. S.D (s)

SAMPLING FLUCTUATION :-

statistics varies from sample to sample, this variation in the value of statistics is called sampling fluctuation

POPULATION MEAN :-

If $x_1, x_2, \ldots x_n$ represent a population of size 'n', then population mean is defined by $\mu = \frac{\sum_{i=1}^{n} x_i}{n}$

POPULATION VARIANCE :-

If $x_1, x_2 \ldots\ldots x_n$ represents population of size 'n' then population variance is defined by:

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{N} \quad (or) \quad E(x^2) - \mu^2$$

CENTRAL LIMIT THEOREM :-

If $\bar{x}$ with mean of a sample size 'n' drawn from a population with mean 'μ' and standard deviation 'σ' then the standard sample mean

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

STANDARD ERROR :-

It is the standard deviation of the sampling distribution. (S.E)

1. S.E of ($\bar{x}$) sample mean $= \sigma/\sqrt{n}$
2. S.E of sample proportion $p = \sqrt{\frac{PQ}{n}}$

   where $Q = 1-P$
3. S.E of two sample means $\bar{x}_1$ & $\bar{x}_2$ is

   i.e $(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}$

4. S.E of 2 proportions $(p_1 - p_2) = \sqrt{\frac{P_1 Q_1}{n_2} + \frac{P_2 Q_2}{n_2}}$

5. S.E of

For a finite population of size 'N' when sample is drawn without replacement.

(1). S.E of sample mean $\bar{x} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$

(2) S.E of sample proportion $\bar{x} = \sqrt{\frac{PQ}{n}} \sqrt{\frac{N-n}{N-1}}$

NOTE:- $\sqrt{\frac{N-n}{N-1}}$ is called correction factor.

1. A population consists of 5 Numbers $\{2, 3, 6, 8, 11\}$ consider all possible samples of size '2' which can be drawn

(1) with replacement from the population

(2) without replacement from the population then find (a) Mean of population

(b) S.D of population

(c) Mean of the sampling distribution of means.

(d) S.D of the sampling distribution of means.

Sol:- (a) Mean of population

$$\mu = \sum_{i=1}^{n} \frac{x_i}{n}$$

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

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

$$\boxed{\mu = 6}$$

(ii) Standard deviation of population :-

$$\sigma^2 = \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{n}$$

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

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

$$\boxed{\sigma^2 = 10.8}$$

$$S.D = \boxed{\sigma = 3.29}$$

1. Drawing samples of size 2 with replacement.

population size N ; Sample size = N^2

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

Means :-

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

(c) Mean of sampling distribution means :-

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$$

$$= [2 + 2.5 + 4 + 5 + 6.5 + 2.5 + 3 + 4.5 + 5.5 + 7 + 4 + 4.5 + 6 + 7 + 8.5 + 5 + 5.5 + 7 + 8 + 9.5 + 6.5 + 7 + 8.5 + 9.5 + 11] / 25$$

$$\boxed{\bar{x} = 6}$$

(d) S.D of sample distributions of means

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

$$= \Big([2-6]^2 + (2.5-6)^2 + (4-6)^2 + (5-6)^2 + (6.5-6)^2 + (2.5-6)^2 + (3-6)^2 + (4.5-6)^2 + (5.5-6)^2 + (7-6)^2 + (4-6)^2 + (4.5-6)^2 + (6-6)^2 + (7-6)^2 + (8.5-6)^2 + (5-6)^2 + (5.5-6)^2 + (7-6)^2 + (8-6)^2 + (9.5-6)^2 + (6.5-6)^2 + (7-6)^2 + (8.5-6)^2 + (9.5-6)^2 + (11-6)^2 \Big] \Big/ 25$$

$$= \Big(16 + 12.25 + 4 + 1 + 0.25 + 12.25 + 9 + 2.25 + 0.25 + 1 + 4 + 2.25 + 0 + 1 + 6.25 + 4 + 12.25 + 0.25 + 1 + 6.25 + 12.25 + 25 + 1 + 0.25 + 1 \Big] \Big/ 25$$

$$\boxed{\sigma_{\bar{x}}^2 = 5.4}$$

$$S.D = \sigma / \sqrt{n} = \sigma = 2.32.$$

(ii) without replacement.

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

$$\text{Mean} = \left\{ \begin{array}{llll} 2.5 & 4 & 5 & 6.5 \\ 4.5 & 5.5 & 7 & \\ 7 & 8.5 & & \\ 9.5 & & & \end{array} \right.$$

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}.$$

$$= \frac{2.5+4+5+6.5+4.5+5.5+7+7+8.5+9.5}{10.}$$

$$\boxed{\bar{x} = 6}$$

S.D. of sample distributions of means.

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

$$= [(2.5-6)^2 + (4-6)^2 + (5-6)^2 + (6.5-6)^2 + (4.5-6)^2 + (5.5-6)^2 + (7-6)^2 + (7-6)^2 + (8.5-6)^2 + (9.5-6)^2]/10$$

$$= [12.25 + 4 + 1 + 0.25 + 2.25 + 0.25 + 1 + 1 + 6.25 + 12.25]/10$$

$$= 40.5/10$$

$$\boxed{\sigma_{\bar{x}}^2 = 4.05}$$

$$\boxed{S.D = \sigma_{\bar{x}} = 2.0125}$$

H.W.

2. Sample of size 2 taken from population $\{1, 2, 3, 4, 5, 6\}$

(i) with replacement

(ii) without replacement.

(a) Mean of population

(b) S.D of population

(c) Mean of the sampling distribution of the means

(d) S.D of sampling distribution of the means

(a) Mean of population :-

$$\mu = \sum_{i=1}^{n} \frac{f_i}{n}$$

$$= \frac{1+2+3+4+5+6}{6}$$

$$\boxed{\mu = 3.5}$$

(b) S.D of population :-

$$\text{Variance} = \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{n}$$

$$= \frac{(1-3.5)^2+(2-3.5)^2+(3-3.5)^2+(4-3.5)^2+(5-3.5)^2 + (6-3.5)^2}{6}$$

$$= \frac{6.25+2.25+0.25+0.25+2.25+6.25}{6}$$

$$\boxed{\sigma^2 = 2.875}$$

$$S.D = \sigma = \sqrt{\text{variance}} = \sqrt{2.875}$$

$$\boxed{\sigma = 1.6957}$$

i. Drawing samples of size 2 with replacement

$= N^n = 6^2 = 36.$

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

Mean of Samples.

$$\begin{bmatrix} 1 & 1.5 & 2 & 2.5 & 3 & 3.5 \\ 1.5 & 2 & 2.5 & 3 & 3.5 & 4 \\ 2 & 2.5 & 3 & 3.5 & 4 & 4.5 \\ 2.5 & 3 & 3.5 & 4 & 4.5 & 5 \\ 3 & 3.5 & 4 & 4.5 & 5 & 5.5 \\ 3.5 & 4 & 4.5 & 5 & 5.5 & 6 \end{bmatrix}$$

$$\bar{x} = \frac{1}{36}\Big[1+1.5+2+2.5+3+3.5+1.5+2+2.5+3+3.5+4+ 2+2.5+3+3.5+4+4.5+2.5+3+3.5+4+4.5+5+ 3+3.5+4+4.5+5+5.5+3.5+4+4.5+5+5.5+6\Big]$$

$$\boxed{\bar{x} = 3.5}$$

S.D of samples = variance $= \dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$

variance $\sigma_{\bar{x}}^2 = \frac{1}{36}\Big[(1-3.5)^2+(1.5-3.5)^2+(2-3.5)^2+(2.5-3.5)^2+(3-3.5)^2+(3.5-3.5)^2 + (2.5-3.5)^2+(3-3.5)^2+(3.5-3.5)^2 + (1.5-3.5)^2+(2-3.5)^2+(4-3.5)^2+(2-3.5)^2+(2.5-3.5)^2+(3-3.5)^2 + (3.5-3.5)^2+(4-3.5)^2+(4.5-3.5)^2+(2.5-3.5)^2+(3-3.5)^2 + (3.5-3.5)^2+(4-3.5)^2+(4.5-3.5)^2+(5-3.5)^2+(3-3.5)^2 + (3.5-3.5)^2+(4-3.5)^2+(4.5-3.5)^2+(5-3.5)^2+(5.5-3.5)^2 + (3.5-3.5)^2+(4-3.5)^2+(4.5-3.5)^2+(5-3.5)^2+(5.5-3.5)^2 + (6-3.5)^2\Big]$

$$= \frac{1}{36}\left[\frac{52.5}{\cancel{18}}\right] = 1.4583$$

S.D $\sigma = \sqrt{\text{variance}} = \sqrt{1.4583} = 1.2076$

2. Drawing Samples of size 2 without replacement

$N_{C_n} = {}^6C_2 = 15$

$$\left\{\begin{matrix} (1,2), (1,3), (1,4), (1,5), (1,6) \\ (2,3), (2,4), (2,5)\ (2,6) \\ (3,4)\ (3,5)\ (3,6) \\ (4,5)\ (4,6) \\ (5,6) \end{matrix}\right\}$$

Mean of samples :-

$$\bar{x} = \frac{1}{15}\Big[(1.5)+(2)+(2.5)+(3)+(3.5)+(2.5)+(3)+(3.5) + (4)+(3.5)+(4)+(4.5)+(4.5)+(5)+(5.5)\Big]$$

$= 52.5$

$\bar{x} = 3.5$

S.D of samples :-

Variance $\sigma_{\bar{x}}^2 = \frac{1}{15}\Big[(1.5-3.5)^2+(2-3.5)^2+(2.5-3.5)^2+(3-3.5)^2+(3.5-3.5)^2 + (2.5-3.5)^2+(3-3.5)^2+(4-3.5)^2+(3.5-3.5)^2+ (4-3.5)^2+(4.5-3.5)^2+(4.5-3.5)^2+(5-3.5)^2+ (5.5-3.5)^2\Big]$

$= \frac{1}{15}\big[17.5\big]$

$\boxed{\sigma_{\bar{x}}^2 = 1.167}$

S.D $= \sigma = \sqrt{\text{variance}} = \sqrt{1.167}$

$= 1.080$

3. The mean height of students in a college is 155 cm. and S.D is 15. wt is the probability than mean height of 36 students is less than 157cm

Sol:- Given that

Mean ht. of students $\mu = 155$ cm

S.D of ht. of students $\sigma = 15$

Mean ht of 36 students $\bar{x} = 157$ cm.

$n = 36$.

$p(\bar{x} \leq 157)$

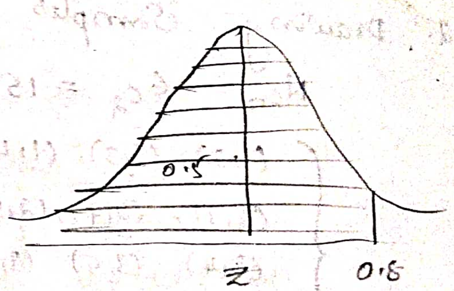
$$Z = \frac{\bar{x}_1 - \mu}{\sigma/\sqrt{n}} = \frac{157-155}{15/\sqrt{36}} = 0.8.$$

$p(\bar{x} < 157) = p(z \leq 0.8)$

$= 0.5 + A(0.8)$

$= 0.5 + 0.2881$

$\boxed{p(\bar{x} < 157) = 0.7881}$



4. A Random sample of size 100 is taken from an infinite population having the mean 76, variance 256 wt is the probability that $\bar{x}$ will be b/w 75 and 78

Sol: $n = 100$

Mean of infinte population $\mu = 76$.

variance, $\sigma^2 = 256$

$S.D = \sigma = 16$

$p(75 \le x \le 78)$

$$z_1 = \frac{75-76}{16/\sqrt{100}} = -0.63$$

$$z_2 = \frac{78-76}{16/\sqrt{100}} = 1.25$$

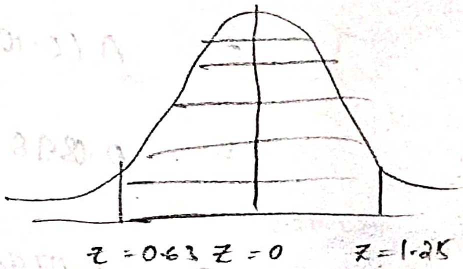
$p(75 \le x \le 78) = p(-0.63 \le z \le 1.25)$

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

$= A(0.63) + A(1.25)$

$= 0.2324 + 0.3944$

$= 0.6268$



5. A random sample of size 64 is taken from a normal population with $\mu = 51.4$ and $\sigma = 68$. What is the probability that the mean of the sample will

(a) exceed 52.9 (b) fall b/w 50.5 and 52.3.

(c) be less than 50.6.

Sol:- $n = 64$.

Mean $= \mu = 51.4$

S.D $= \sigma = 68$.

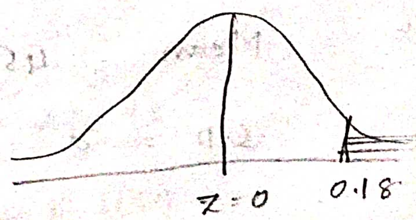
(a) $P(x \ge 52.9)$

$$z_1 = \frac{52.9-51.4}{68/\sqrt{64}} = 0.18$$

$P(x \ge 0.18) = 0.5 - A(0.18)$

$= 0.5 - 0.0675$

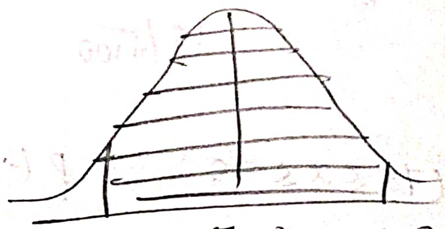
$= 0.4325$



(b) $P(50.5 \leq x \leq 52.3)$.

$$Z_1 = \frac{50.5 - 51.4}{68/\sqrt{64}} = -0.10.$$

$$Z_2 = \frac{52.3 - 51.4}{68/\sqrt{64}} = 0.10.$$



$P(-0.10 \leq Z \leq 0.10)$

$= A(0.10) + A(0.10)$

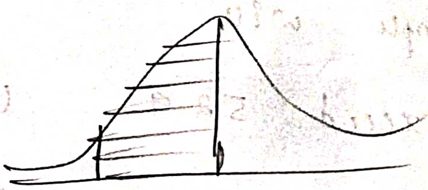
$= 0.0398 + 0.0398$

$= 0.0796.$

(c) be less than 50.6. $P(Z \leq 50.6)$

$$Z_1 = \frac{50.6 - 51.4}{68/\sqrt{64}} = -0.09.$$

$P(Z \leq -0.09) = 0.5 - A(0.09)$



$= 0.5 - 0.0359$

$= 0.4641$

6. A random sample of size 64 is taken from an infinite population having the mean 45 and SD-8. What is the probability that x will be b/w 46 and 47.5

Sol:- Mean = 45 $\qquad n = 64$

S.D = 8

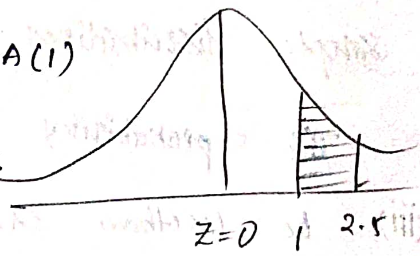
$p(46 \le x \le 47.5)$

$$z_1 = \frac{46-48}{8/\sqrt{64}} = 1$$

$$z_2 = \frac{47.5-45}{8/\sqrt{64}} = 2.5$$

$p(1 \le x \le 2.5) = \cancel{0.8} - A(2.5) - A(1)$

$= 0.4938 - 0.3413$



$= 0.1525$

(7). A random sample of size 81 is taken from a population having mean 65 and S.D 10. Then probability of ~~sample~~ '$\bar{x}$' will be

(i) b/w 66 and 68 (iv) exceed 58

(ii) exceed 80 (v) Less than 49.

(iii) be lessthan 32

Sol: Given

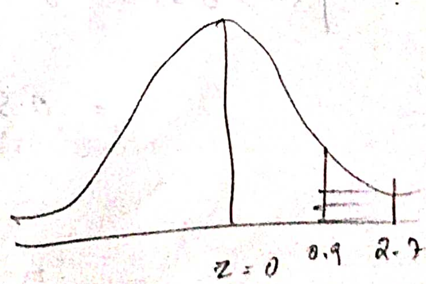
$n = 81$

$\mu = 65$

$\sigma = 10.$

(i) $p(66 \le x \le 68)$

$$z_1 = \frac{66-65}{10/\sqrt{81}} = 0.9$$



$$z_2 = \frac{68-65}{10/9} = 2.7.$$

$p(0.9 \le z \le 2.7) = A(2.7) - A(0.9)$

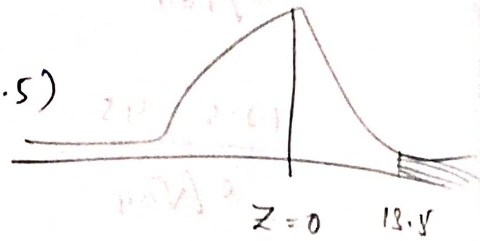
$= 0.4965 - 0.3159$

$= 0.1806$

(ii) $p(x \geq 80)$

$$Z = \frac{80-65}{10/\sqrt{81}} = 13.5$$

$$p(Z \geq 13.5) = 0.5 - A(13.5)$$



$$= 0.5 -$$

~~X. Sample distributions of means (σ is known)~~

~~The probability distribution of $\bar{x}$ is called~~

(iii) be lessthan 32.

$p(x \leq 32)$

$$Z = \frac{32-65}{\frac{10}{\sqrt{81}}} = -29.7$$

$p(x \leq 29.7) =$

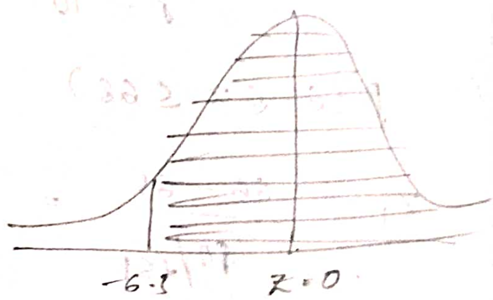
(iv) exceed 58.

$p(x \geq 58)$

$$Z = \frac{58-65}{\frac{10}{\sqrt{81}}} = -6.3.$$

$$p(x \geq 6.3) = 0.5 + A(6.3)$$

$$= 0.5 +$$

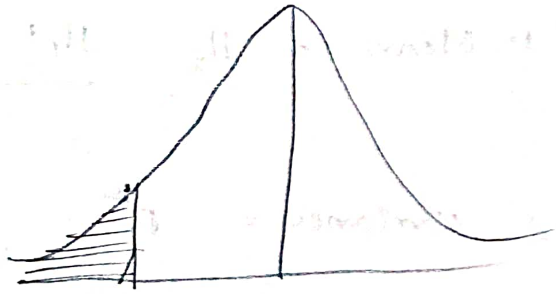


(iv) less than 49.

$p(x \le 49)$

$$z = \frac{49-65}{\frac{10}{\sqrt{81}}} = -14.4$$

$p(z \le 49) = 0.5 - A(-14.4).$



Sampling distribution of means (σ is Known):-

The probability distribution of $\bar{x}$ is called sampling distribution of means

INFINITE POPULATION:-

Suppose the samples are drawn ~~wi~~ from an infinite population or sampling is done with replacement then

1. Mean of sampling distribution is

$$\mu_{\bar{x}} = \frac{\mu_1 + \mu_2 + \cdots \mu_n}{n}$$

2. The variance of sampling distribution

$$\sigma_{\bar{x}}^2 = \frac{\sigma_1^2 + \sigma_2^2 + \cdots \sigma_n^2}{n} = \frac{\sigma^2}{n}.$$

3.

FINITE POPULATION:-

Consider a finite population of 'N' with mean 'μ'. S.D 'σ'. draw all possible samples of size n without replacement from the population

1. Mean $= \mu_{\bar{x}} = \dfrac{\mu_1 + \mu_2 + \cdots \mu_n}{n}$

2. Variance $= \sigma_x^2 = \dfrac{\sigma}{\sqrt{n}} \left[\sqrt{\dfrac{N-n}{N-1}}\right]$

Where $\left[\dfrac{N-n}{N-1}\right]$ is called correction factor.

NOTE :-

Let population size 'N' and sample size 'n', then

1. Total no. of samples with replacement is N^n

2. If total no. of samples without replacement is N_{C_n}

1. Find S.E of a sample size 64 is taken from a normal population with variance '8'.

Sol:- S.E of sample size is $\dfrac{\sigma}{\sqrt{n}} =$

$S.D = \sqrt{\text{variance}}$

$= \sqrt{8} = 2.828$

$S.E = \dfrac{\sigma}{\sqrt{n}} = \dfrac{2.83}{\sqrt{64}} = \dfrac{2.83}{8} = 0.3535$

Determine the population correction factor for a population size of 12 and sample size of 4.

Sol: correction factor is $= \frac{N-n}{N-1}$

population size (N) = 12

sample size $n = 4$

correction factor is $= \frac{12-4}{12-1} = \frac{8}{11}$

$= 0.7272.$

3. If the population 16, 14, 12, 8, 24, 20 by drawing samples of size '3' without replacement. Determine (i) population mean. (ii) population S.D (iii) Mean of sample distribution means (iv) S.D of sample distributions means.

Sol: Given population is

$\{8, 12, 14, 16, 20, 24\}$

population size $N = 6$

Sample size $n = 3$

No. of elements in sample without replacement

$^{N}C_{n} = {}^{6}C_{3} = 20$

(i) population mean $(\mu) = \frac{\sum_{i=1}^{n} f_i}{n}$

$= \frac{8+12+14+16+20+24}{6}$

$= 15.67.$

(ii) S.D of population σ =

$$\text{variance} = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}$$

$$= \frac{(8-15.7)^2 + (12-15.7)^2 + (14-15.7)^2 + (16-15.7)^2 + (20-15.7)^2 + (24-15.7)^2}{6}$$

$$= \frac{163.34}{6}$$

$$= 27.22$$

$$S.D = \sqrt{\text{variance}}$$

$$= \sqrt{27.22}$$

$$\boxed{S.D(\sigma) = 5.22}$$

(iii) Population is $\{8, 12, 14, 16, 20, 24\}$

Samples of sizes '3' are

(8, 12, 14), (8, 12, 16), (8, 12, 20), (8, 12, 24)

(8, 14, 16), (8, 14, 20), (8, 14, 24), (8, 16, 20)

(8, 16, 24), (8, 20, 24),

(12, 14, 16), (12, 14, 20) (12, 14, 24) (12, 16, 20)

(12, 16, 24), (12, 20)(24)

(14, 16, 20), (14, 16, 24), (14, 20, 24)

(16, 20, 24)

Means of samples.

11.33, 12, 13.33, 14.7
16, 14, 15.33, 14.7
16., 15.33, 16.7, 16.00
17.33, 18.7.

Means of samples.

$$\begin{bmatrix} 11.33 & 12 & 13.33 & 14.7 \\ 12.7 & 14 & 15.33 & 14.7 \\ 16 & 17.33 & 14 & 15.33 \\ 16.7 & 16 & 17.3 & 18.7 \\ 16.7 & 18 & 19.33 & 20. \end{bmatrix}$$

$$\bar{x} = \mu_{\bar{x}} = \frac{1}{20}\left[11.33 + 12 + 13.33 + 14.7 + 12.7 + 14 + 15.33 + 14.7 + 16 + 17.33 + 14 + 15.33 + 16.7 + 16 + 17.3 + 16.7 + 18 + 19.33 + 20 \right]$$

$$= \frac{313.18}{20} = 15.66 = 15.7$$

$$\text{Variance} = \frac{1}{20}\left[19.10 + 5.62 + 13.69 + 1 + 9 + 2.89 + 0.1369 + 1 + 0.09 + 2.6569 + 2.89 + 0.1369 + 4 + 0.09 + 2.56 + 9 + 4 + 5.29 + 13.1769 + 18.49 \right]$$

$$= \frac{114.8136}{20}$$

$$\sigma_{\bar{x}}^2 = 5.74$$

$$S.D = \sigma = 2.40.$$

4. Given that sample size is 36. No. of samples is 300 population size is 1500. population S.D is 0.48, population mean $\mu = 22.4$. Determine the expected no. of samples having their means b/w (i) 22.39 and 22.41 (ii) ≥ 22.42 (iii) < 22.37 (iv) > 22.41

Sol:- Given that: $n = 36$ $\quad N_s = 300$

$N = 1500$ $\quad \sigma = 0.48$ $\quad \mu = 22.4$

$\therefore$ Standard Normal mean

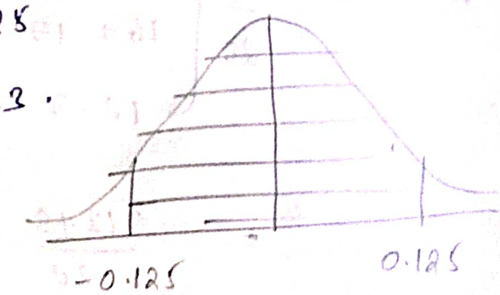
$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

(i) No. of samples mean b/w 22.39 & 22.41

$P(22.39 \leq x \leq 22.41)$

$$z_1 = \frac{22.39 - 22.4}{0.48/\sqrt{36}} = -0.125$$
$$= -0.13$$

$$z_2 = \frac{22.41 - 22.4}{0.48/\sqrt{36}} = 0.125$$
$$= 0.13.$$



$P(-0.125 \leq z \leq 0.125)$

$= P(-0.13 \leq z \leq 0.13)$

$= A(0.13) + A(0.13)$

$= 0.0517 + 0.0517$

$= 0.1034$

No. of samples b/w 22.39 & 22.41

$= 300 \times 0.1034$

$= 31$

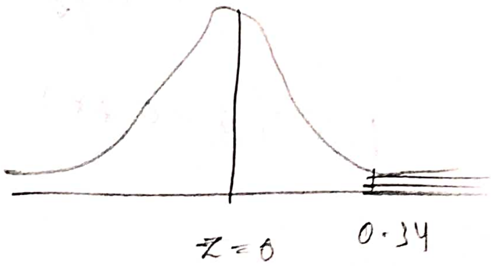
$p(x \geq 22.42)$

$$z_1 = \frac{22.42 - 22.4}{0.48/\sqrt{36}} = 0.25.$$

$p(z_1 \geq 0.34) = 0.5 - A(0.25)$

$= 0.5 - 0.0987$

$= 0.4013$



No. of samples $= N_s \times p(x \geq 22.42.$

$= 300 \times 0.4013$

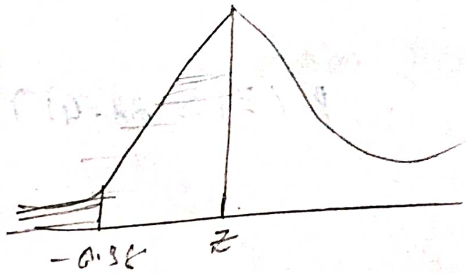
$= 120.4.$

(ii) $p(x \leq 22.37)$

$$z_1 = \frac{22.37 - 22.4}{0.48/\sqrt{36}} = -0.38$$

$p(z_1 \leq -0.38) = 0.5 - 0.1480$

$= 0.352$



No. of Samples $= N_s \times (p(z \leq -0.38)$

$= 300 \times 0.352$

$= 105.6$

$= 106.$

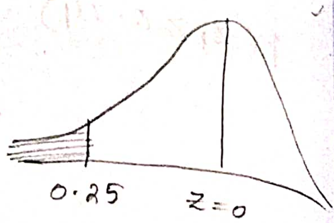
(iv) less than 22.38

$p\ (x \leq 22.38)$

$$z_1 = \frac{22.38 - 22.4}{0.48/\sqrt{6}} = -0.25$$

$p\ (z_1 \leq -0.25) = 0.5 - 0.0987$

$= 0.4013 \times 300.$



No. of samples $= N_s \times p(z_1 \leq 0.25)$

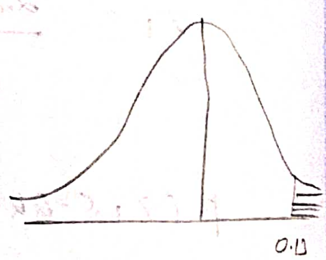
$= 120.4$

(v). More than 22.41

$p\ (x \geq 22.41)$

$$z_1 = \frac{22.41 - 22.4}{0.48/\sqrt{6}} = 0.125$$

$= 0.13.$



$p\ (z_1 \geq 22.41) = 0.5 - A\ (0.13)$

$= 0.5 - 0.0517$

$= 0.4483$

No. of samples.

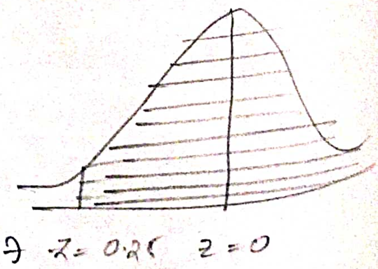
$= 134.5$

$= 135.$

(vi) more 22.38

$p\ (x \geq 22.38)$

$$z_1 = \frac{22.38 - 22.4}{0.48/\sqrt{6}} = -0.25$$



$p(z_1 \geq \cancel{22}.25) = 0.5 + 0.0987$

$= 0.5987$

No. of samples $= 300 \times 0.5987$

$= 179.61$

$= 180.$

(ii) less than 22.41

$p(x \leq 22.41)$

$$z_1 = \frac{22.41 - 22.4}{0.48/6} = 0.125 \approx 0.13.$$

$p(z_1 \leq 0.13) = 0.5 + A(0.13)$

$= 0.5 + 0.0517$

$= 0.5517$

No. of samples $= 300 \times 0.5517$

$= 165.51$

$= 166.$

ESTIMATION :-

ESTIMATOR :-

Procedure to find an unknown population parameter is called an estimator.

point estimator :-

If an estimate of a population parameter is gives a single value, then, the estimator is called point estimator

Eg:- 1. '$\bar{x}$' is point estimator of population mean 'μ'

2. s^2 is point estimator of population variance 'σ^2'

3. s is point estimator of population S.D 'σ'

Unbiased estimator :-

Let $\hat{\theta}$ be an estimator of 'θ' and, the statistic $\hat{\theta}$ is said to be unbiased estimator of 'θ' if $\hat{\theta} = \theta$, Otherwise it is called biased estimator $(\hat{\theta} \neq \theta)$

* Variance of a point estimator :-

If $\hat{\theta}_1$, $\hat{\theta}_2$ are two unbiased estimators of 'θ' of the same population, If $\sigma^2_{\hat{\theta}_1} < \sigma^2_{\hat{\theta}_2}$, then we can say $\hat{\theta}_1$ is more efficient than $\hat{\theta}_2$.

Interval estimator :-

An interval estimator of a population parameter 'θ' is an interval is of the form $(\hat{\theta}_L < \theta < \hat{\theta}_U)$

where $\hat{\theta}_L$ is lower limit and

$\hat{\theta}_U$ is upper limit.

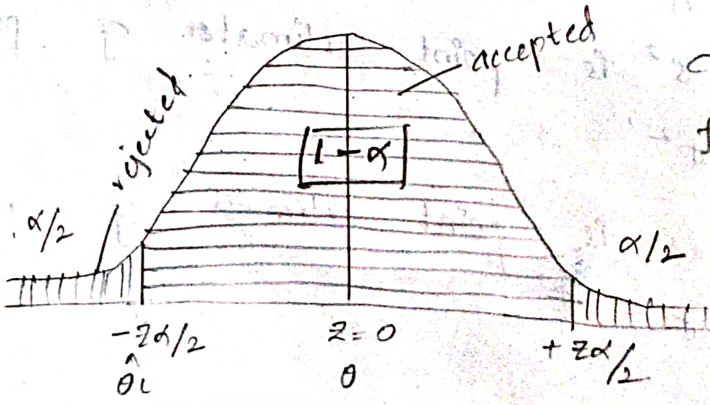
* CONFIDENCE INTERVAL :-

$\hat{\theta}_L$ and $\hat{\theta}_U$ are ~~upper~~ lower and upper confidence limits of 'θ' such that $P(\hat{\theta}_L < \theta < \hat{\theta}_U) = (1-\alpha)$, where $0<\alpha<1$, then the interval $(\hat{\theta}_L < \theta < \hat{\theta}_U)$ is called $(1-\alpha) \times 100\%$ confidence interval

Eg:- 1. If $\alpha = 0.02$, we have 98% confidence interval

2. If $\alpha = 0.05$, we have 95% confindence interval

3. If $\alpha = 0.01$, we have 99% confidence interval



Confidence interval for population mean.

* Confidence interval for population mean :-

Let $\bar{x}$ is sample mean from the population having mean 'μ' and S.D 'σ' where sample size is 'n'

We know that

standard sample mean

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

Let $\hat{\theta}_L$ and $\hat{\theta}_U$ are interval estimators of 'θ',

$\therefore$ $p(\hat{\theta}_L < \theta < \hat{\theta}_U) = 1-\alpha$.

$$P\left(-Z_{\alpha/2} < Z < Z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(-Z_{\alpha/2} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < Z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{x}-\mu < Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$P\left(Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} > -\bar{x}+\mu > -Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$P\left(x + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} > \mu > x - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$P\left(x - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < x + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha.$$

$$\hat{\theta}_L < \theta < \hat{\theta}_U$$

$\therefore$ $(1-\alpha)\times 100\%$ confidence interval for population mean is

$$\left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

Maximum error (E) for Large samples :-
($n \geq 30$)

The sample mean estimate $\bar{x}$, very rarely equals to population mean

$\therefore$ The maximum error is given by

$$|\bar{x} - \mu| = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

i.e $\boxed{E = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}.}$

NOTE :-

To find sample size by using maximum error

We know that

$$E = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

$$\Rightarrow \sqrt{n} = Z_{\alpha/2} \frac{\sigma}{E}$$

$$\boxed{n = \left[Z_{\alpha/2} \frac{\sigma}{E}\right]^2}$$

where n = Sample size

1. At an electronic company considered 40 technician if the performance of getting task a mean of 12.73 min and S.D 2.06, Consider sample mean and population mean.

(i) Wt can we say with 99% confidence about maximum error

(ii) Construct 98% confidence interval

(iii) With what confidence we can assert that the sample mean does not differ from true mean by 30 sec

Sol:- Given that sample size $n = 40$

population mean $\mu = 12.73$ mins

Standard deviation $S.D = \sigma = 2.06$ mins

and given

sample mean $\bar{x} = 12.73$.

(ii) finding max error for 99% confidence

Sample space $(n) = 40 \geq 30 \longrightarrow$ large

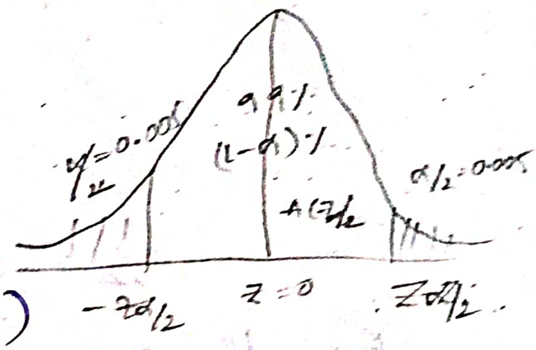
Maximum error $(E) = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$(1-\alpha) \times 100$ confidence $= 99\%$

$(1-\alpha) = 0.99$

$\alpha = 1 - 0.99 = 0.01$

$\alpha/2 = 0.005$



By diagram $\alpha/2 = 0.5 - A(Z_{\alpha/2})$

$A(Z_{\alpha/2}) = 0.5 - \alpha/2$.

$A(Z_{\alpha/2}) = 0.5 - 0.005$

$A(Z_{\alpha/2}) = 0.495$

$A(2.58) = 0.495$

$\boxed{Z_{\alpha/2} = 2.58}$

Maximum error $(E) = 2.58 \times \frac{2.06}{\sqrt{40}}$

$= 0.84$.

(ii) Construct 98% confidence interval.

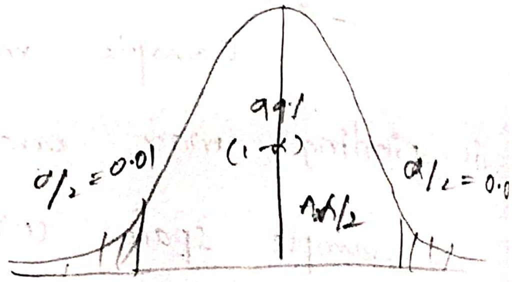
$(1-\alpha)100\%$ confidence interval for population mean

$$\bar{x} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Given $\bar{x} = 12.73$

$\sigma = 2.06$

$n = 40 \geq 30$



$(1-\alpha)100 = 98\%$

$(1-\alpha)\cancel{100} = 0.98 \times \cancel{100}$

$1-\alpha = 0.98$

$\alpha = 1 - 0.98$

$\alpha = 0.02$

$\alpha/2 = 0.01$

By diagram $\alpha/2 = 0.5 - \cancel{area}\, A(Z_{\alpha/2})$

$A(Z_{\alpha/2}) = 0.49$

$A(2.33) = 0.49$

$$\boxed{Z_{\alpha/2} = 2.33}$$

98% confidence interval is

$$12.73 - 2.33\left[\frac{2.06}{\sqrt{40}}\right] \leq \mu \leq 12.73 + 2.33\left(\frac{2.06}{\sqrt{40}}\right)$$

$$11.9710 \leq \mu \leq 13.4889$$

(iii) find the confidence

$E = 30$ sec

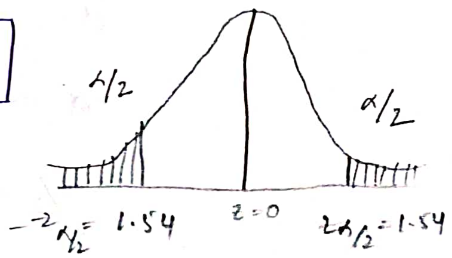
$E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$E = |\bar{x} - \mu| = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$30 \sec \times \frac{1}{60} = Z_{\alpha/2} \cdot \frac{2.06 \times 60}{\sqrt{40}}$

$Z_{\alpha/2} = 1.54$

$\boxed{A(Z_{\alpha/2}) = 0.4382}$



By diagram.

$\alpha/2 = 0.5 - A(Z_{\alpha/2})$

$= 0.5 - 0.4382$

$\alpha/2 = 0.0618$

$\boxed{\alpha = 0.1236}$

$\therefore (1-\alpha)\, 100\% = 0.8764$

confidence $= (1-\alpha)\, 100\%$.

$= 87.64\%$.

1. The diff b/w sample mean and true mean is not more than ~~87.64~~ 30 sec with 87.64% confidence.

2. $\sigma = 1.40$ mins. Determine how large should be sample size if it is with 95% confidence that the max. error is atmost 0.25

Sol:- Given that

population S.D $\sigma = 1.40$ min

$E = 0.25$

$(1-\alpha)\, 100\% = 95\%$.

$(1-\alpha)\, \cancel{100} = 0.95 \times \cancel{100}$

$(1-\alpha) = 0.95$

We know that max error

$E = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$.

$$\sqrt{n} = Z_{\alpha/2}\frac{\sigma}{E}$$

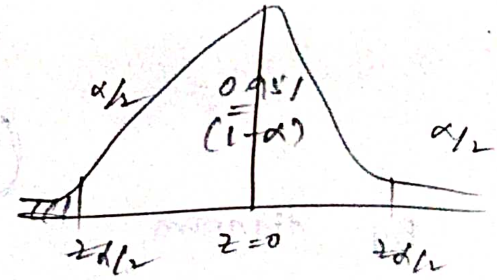
$$n = \left[Z_{\alpha/2}\frac{\sigma}{E}\right]^2$$

$(1-\alpha) = 0.95$

$\alpha = 1 - 0.95$

$\alpha = 0.05$

$\alpha/2 = 0.025$



$\alpha/2 = 0.5 - A(Z_{\alpha/2})$

$A(Z_{\alpha/2}) = 0.5 - 0.025$

$A(Z_{\alpha/2}) = 0.475$

$A(1.96) = 0.475$

$Z(\alpha/2) = 1.96$

$$n = \left(1.96 \times \frac{1.40}{0.25}\right)^2$$

$$\boxed{n = 120.4736}$$

3. A random sample of size 100 has a S.D of 5. What can you say about the maximum error with 95% confidence

Soln- $n = 100$

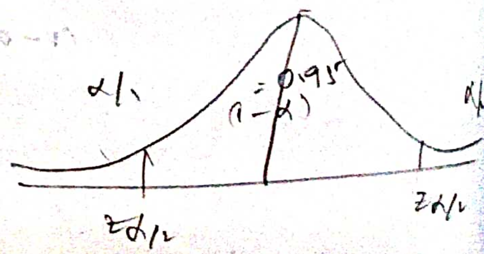
$\sigma = 5.$

$(1-\alpha)\,100 = 95\%.$

$(1-\alpha)\,\cancel{100} = 0.95 \times \cancel{100}$

$1-\alpha = 0.95$

$\boxed{\alpha = 0.05}$



$\alpha/2 = 0.025$

$\alpha/2 = 0.5 - A(Z_{\alpha/2})$

$A(Z_{\alpha/2}) = 0.5 - 0.025$

$= 0.475$

$\boxed{Z_{\alpha/2} = 1.96}$

$E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$= 1.96 \cdot \frac{5}{\sqrt{100}}$

$\boxed{E = 0.98}$

7. A random sample of size n 81 was taken whose variance is 20.25 and mean is 32 construct 98% confidence interval

Sol:- $n = 81$ $\bar{x} = 32.$

$\sigma^2 = 20.25$

$\sigma = 4.5.$

$\mu = 32.$

$(1-\alpha)100 = 98\%.$

$(1-\alpha)\cancel{100} = 0.98 \times \cancel{100}$

$\alpha = 1 - 0.98$

$\alpha = 0.02$

$\frac{\alpha}{2} = 0.01$

$\alpha/2 = 0.5 - A(Z_{\alpha/2})$

$= 0.5 - 0.01$

$A(Z_{\alpha/2}) = 0.49$

$\boxed{Z_{\alpha/2} = 2.33}$

Confidence interval

$$x - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq x + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$32 - 2.33 \frac{4.5}{\sqrt{81}} \leq \mu \leq 32 + 2.33 \frac{4.5}{\sqrt{81}}$$

$$30.835 \leq \mu \leq 33.165.$$

5. A random sample of 100 such teachers revealed mean salary of Rs 287 with S.D of Rs. 48 and find the degree of confidence

Incomplete question

Sol:- $n = 100$

$\bar{x} = 287/-$

$\sigma = 48/-$

$$|287 - 287| = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$|287 - 287| = z_{\alpha/2} \cdot \frac{48}{\sqrt{100}}.$$

$$z_{\alpha/2} = 0.206$$

$$\alpha/2 = 0.5 - A(z_{\alpha/2})$$

$$= 0.5 - 0.208$$

$$\alpha/2 = 0.292$$

$$\boxed{\alpha = 0.584} \approx 0.6$$

degree of confidence

$= (1-\alpha)\, 100\%$

$= (1-0.6)\, 100\%$

$= (0.4)\, 100\%$

$= 40\%$