S. No.	Date	Title	Page No.	Teacher's Sign / Remarks

17/09/2024

MODULE - 1

PROBABILITY DISTRIBUTIONS

→ **RANDOM VARIABLE** :- In a random experiment, if a real variable is associated with every outcome, then it is called random variable.

OR

A random variable is a function that assigns a real number to every sample point in the sample space of a random experiment.

→ Random variable is denoted by 'x' 'y' 'Z' etc.

Example :-

① Tossing a coin and observe the outcome

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

x = no. of heads turning up

$x = 1, 0$

$\boxed{x = 0, 1}$

② Tossing a coin two times & observe the outcome

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

x = no. of heads turning up

$x = 2, 1, 1, 0$

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

• DISCRETE RANDOM VARIABLE :-

If a random variable takes finite (or) countabaly infinite no. of values then it is called discrete random variables.

Example :-

1. Tossing a coin & observing the outcome
2. Tossing a coin & observing the no. of heads turning up.
3. Throwing a dice & observing the no. of faces.

• PROBABILITY MASS FUNCTION

[Discrete Probability Distributions]

If 'x' is a discrete random variable i.e., x_1, x_2, x_3 then p is called probability $p(x_1), p(x_2), \ldots p(x_n)$, then p is called probability mass function (pmf) of 'x' provided if the following conditions are satisfied.

(i) $p(x) \geq 0$

(ii) $\Sigma p(x) = 1$

• NOTE :-

Mean$(\mu) = \Sigma x . p(x)$

Variance $(v) = [\Sigma x^2 . p(x)] - (\mu)^2$

Standard Deviation $(\sigma) = \sqrt{v}$

TYPE 1 PROBLEMS :-

1. A coin is tossed twice, a random variable x represents the no. of heads turning up. Find the discrete probability distribution foo capital x. Also find Mean, variance & std deviation.

→ A coin is tossed twice

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

given :- Random variable "x" represents no. of heads turning up

$x : 2, 1, 1, 0$

$x = 0, 1, 2$

- $p(x=0$, i.e., no head$) = 1/4$
- $p(x=1$, i.e., one head$) = 1/4 + 1/4 = 1/2$
- $p(x=2$, i.e., two head$) = 1/4$

→ We observe that,

(i) $p(x) \geq 0$

(ii) $\Sigma p(x) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$

∴ $p(x)$ is probability mass function.

x	0	1	2
$p(x)$	1/4	1/2	1/4

[Discrete probability distribution]

Mean $(\mu) = \Sigma x . p(x)$

$= 0(1/4) + 1(1/2) + 2(1/4)$

$\boxed{\mu = 1}$

Variance (v) = $\left[\Sigma x^2 p(x)\right] - \mu^2$

$= \left[0^2(1/4) + 1^2(1/2) + 2^2(1/4)\right] - 1^2$

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

Std deviation (σ) = $\sqrt{v}$

$= \sqrt{1/2}$

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

---

② When a coin is tossed thrice, then find Mean, Variance, std deviation, x represents no. of heads turning up.

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

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

$x = 0, 1, 2, 3$

$P(x=0) = 1/8$

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

$P(x=2) = 3/8$

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

x	0	1	2	3
$p(x)$	$1/8$	$3/8$	$3/8$	$1/8$

we observe that,

(i) $p(x) \geq 0$,

(ii) $\Sigma p(x) = 1$

∴ probability mass function

Mean (μ): $0(1/8) + 1(3/8) + 2(3/8) + 3(1/8)$

$= \boxed{3/2}$

Variance = $1(3/8) + 4(3/8) + 9(1/8) - \frac{9}{4}$

$= \frac{3}{8} + \frac{12}{8} + \frac{9}{8} - \frac{9}{4}$

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

Std deviation = $\sqrt{3/4} = \boxed{\frac{\sqrt{3}}{2}}$

---

③ A random experiment of tossing a die twice is performed, a random variable x on this sample space is defined to be the sum of 2 numbers, turning up on the toss. Find the discrete probability distribution of the random variable x, & compute the corresponding Mean & std deviation.

→ S = {

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

}

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

$P(X=2) = 1/36$	$P(X=8) = 5/36$
$P(X=3) = 2/36$	$P(X=9) = 4/36$
$P(X=4) = 3/36$	$P(X=10) = 3/36$
$P(X=5) = 4/36$	$P(X=11) = 2/36$
$P(X=6) = 5/36$	$P(X=12) = 1/36$
$P(X=7) = 6/36$	

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

$$= 2\left(\frac{1}{36}\right) + 3\left(\frac{2}{36}\right) + 4\left(\frac{3}{36}\right) + 5\left(\frac{4}{36}\right) + 6\left(5/36\right) + 7\left(6/36\right) + 8\left(5/36\right) + 9\left(4/36\right) + 10\left(3/36\right) + 11\left(2/36\right) + 12\left(1/36\right)$$

$$= \boxed{7}$$

$$\text{Variance } (v) = \left[\Sigma x^2 . p(x)\right] - \mu^2$$

$$= 4(1/36) + 9\left(\frac{2}{36}\right) + 16\left(\frac{3}{36}\right) + 25\left(\frac{4}{36}\right) + 36\left(\frac{5}{36}\right) + 49\left(\frac{6}{36}\right) + 64\left(\frac{5}{36}\right) + 81\left(\frac{4}{36}\right) + 100\left(\frac{3}{36}\right) + 121\left(\frac{2}{36}\right) + 144\left(1/36\right)$$

$$= \boxed{5.75}$$

Std deviation $= \sqrt{v}$

$$= \sqrt{5.75}$$

$$= \boxed{2.398}$$

TYPE 2 :-

④ A random variable x has the following discrete probability distributions for the various values of x.

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

(i) Find k

(ii) Evaluate $p(x<6)$

$p(x \geq 6)$

$p(3 \leq x \leq 6)$ & also find the probability distribution.

→ Given, Discrete Probability Distribution

∴ It must satisfy (i) $p(x) \geq 0$

(ii) $\Sigma p(x) = 1$

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

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

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

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

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

$$\boxed{K=-1} \text{ or } \boxed{K=\frac{1}{10}}$$

x	0	1	2	3	4	5	6	7
$p(x)$	0	1/10	2/10	2/10	3/10	1/100	1/50	17/100

- $p(x<6) = 1/10 + 2/10 + 2/10 + 3/10 + 1/100 = \boxed{81/100}$
- $p(x \geq 6) = \frac{1}{50} + \frac{17}{100} = \boxed{\frac{19}{100}}$
- $p(3 \leq x \leq 6) = 2/10 + 3/10 + 1/100 + 1/50 = \boxed{\frac{33}{100}}$

⑤ Find the value of k such that the following distribution represents a finite probability function, hence find its mean, std deviation & also find probability of $p(x \leq 1)$, $p(x > 1)$, $p(-1 < x \leq 2)$

x	-3	-2	-1	0	1	2	3
$p(x)$	K	2K	3K	4K	3K	2K	K

→ $K + 2K + 3K + 4K + 3K + 2K + K = 1$

$16K = 1$

$\boxed{K = \frac{1}{16}}$

x	-3	-2	-1	0	1	2	3
$p(x)$	1/16	2/16	3/16	4/16	3/16	2/16	1/16

WKT, Mean $\mu = \Sigma x . p(x)$

$= -3(\frac{1}{16}) + (-2)(2/16) + (-1)(3/16) + 0() + 1(3/16) + 2(2/16) + 3(1/16)$

$\boxed{\mu = 0}$

variance $(v) = [\Sigma x^2 p(x)] - \mu^2$

$= (-3)^2(1/16) + (-2)^2(2/16) + (-1)^2(3/16) + 0() + 1^2(3/16) + 4(2/16) + 9(1/16) - 0$

$= \boxed{5/2}$

std deviation $= \boxed{\sqrt{\frac{5}{2}}}$

- $p(x \leq 1) = \frac{1}{16} + \frac{2}{16} + \frac{3}{16} + \frac{4}{16} + \frac{3}{16} = \boxed{\frac{13}{16}}$
- $p(x > 1) = \frac{2}{16} + \frac{1}{16} = \boxed{\frac{3}{16}}$
- $p(-1 < x \leq 2) = \frac{2}{16} + \frac{3}{16} + \frac{4}{16} = \boxed{\frac{9}{16}}$

⑥ A probability distribution of a finite random variable x is given by the following table.

x	-2	-1	0	1	2	3
$p(x)$	0.1	K	0.2	2K	0.3	K

(i) Find K

(ii) Mean, variance & std deviation

→ $4K + 0.6 = 1$

$4K = 0.4$

$\boxed{K = 0.1}$

x	-2	-1	0	1	2	3
$p(x)$	0.1	0.1	0.2	0.2	0.3	0.1

Mean $\mu = \Sigma(x)p(x)$

$= -2(0.1) - 1(0.1) + 0() + 1(0.2) + 2(0.3) + 3(0.1)$

$= \boxed{0.8}$

variance $= 4(0.1) + 0.1 + 0 + 0.2 + 4(0.3) + 9(0.1) - 0.8^2$

$= \boxed{2.16}$

std deviation $= \boxed{\sqrt{2.16}}$

TYPE 3

⑦ A random variable x has $p(x) = 2^{-x}$, $x = 1, 2, 3$, show that $p(x)$ is a probability mass $\underline{fun}$. Also find $p(x\ \text{even})$, $p(x \text{ being divisible by } 3)$ & $p(x \geq 5)$.

→ Given, $p(x) = 2^{-x}$

$p(x) = \frac{1}{2^x}$, $x = 1, 2, 3, \ldots$

• $p(x) = \frac{1}{2^1}, \frac{1}{2^2}, \frac{1}{2^3}$

$\therefore \boxed{p(x) \geq 0}$

• $\Sigma p(x) = \Sigma \frac{1}{2^x}$

$= \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots\ldots$

It looks like $a + ar + ar^2 + \ldots\ldots = \frac{a}{1-r}$

$a = \frac{1}{2}$, $r = \frac{1}{2}$

On substituting, $\boxed{\Sigma p(x) = 1}$

Hence, it satisfied both the condition.

$\therefore$ it is probability mass function.

• $p(x\ \text{even}) = \sum^{\infty} \frac{1}{2^x}$

$x = 2, 4, 6, \ldots$

$= \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \ldots$

$a = \frac{1}{2^2}$, $r = \frac{1}{2^2}$

$= \frac{\frac{1}{2^2}}{1 - \left(\frac{1}{2^2}\right)} = \frac{1/4}{3/4} = \boxed{\frac{1}{3}}$

• $p(x \text{ being divisible by } 3) = \sum^{\infty} \frac{1}{2^x}$

$x = 3, 6, 9, \ldots$

$= \frac{1}{2^3} + \frac{1}{2^6} + \frac{1}{2^9}$

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

$= \frac{\frac{1}{2^3}}{1 - \left(\frac{1}{2^3}\right)} = \frac{\frac{1}{8}}{7/8} = \boxed{1/7}$

• $p(x \geq 5) = \sum^{\infty} \frac{1}{2^x}$

$x = 5, 6, 7, \ldots$

$= \frac{1}{2^5} + \frac{1}{2^6} + \frac{1}{2^7}$

$a = \frac{1}{2^5}$, $r = \frac{\frac{1}{2^6}}{1/2^5} = \frac{1}{2}$

$= \frac{\frac{1}{2^5}}{1 - \frac{1}{2}} = \boxed{\frac{1}{2^4}}$

⑧ If x is a discrete random variable taking values 1, 2, 3 etc. with $p(x) = \frac{1}{2} \cdot \left(\frac{2}{3}\right)^x$, find probability of x being an odd number, by first establishing that $p(x)$ is a probability mass function.

→ Given, $p(x) = \frac{1}{2}\left(\frac{2}{3}\right)^x$, $x = 1, 2, 3, \ldots$

$p(x) = \frac{1}{3}, \frac{2}{9}, \frac{4}{27}, \ldots$

$\therefore \boxed{p(x) \geq 0}$

• $\Sigma p(x) = \Sigma \frac{1}{2}\left(\frac{2}{3}\right)^x$

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

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

$= \left\{\begin{matrix} a + ar + ar^2 + \cdots\cdots \\ 2/3 + 2/3(2/3)^2 + 2/3(2/3)^3 + \cdots\cdots \\ a = 2/3,\ r = 2/3 \end{matrix}\right\}$

$= \frac{1}{2}\left[\frac{a}{1-r}\right]$

$= \frac{1}{2}\left[\frac{2/3}{1-2/3}\right]$

$= \frac{1}{2}\left[\frac{2/3}{1/3}\right]$

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

$= \boxed{1}$

$\Rightarrow \boxed{\Sigma p(x) = 1}$ Both the conditions are satisfied, Hence, given $p(x)$ probability mass func.

- $p(x$ being an odd number$)$

$= \sum_{x=1,3,5,7..}^{\infty} p(x)$

$= \sum_{x=1,3,5,...} \frac{1}{2}\left(\frac{2}{3}\right)^x$

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

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

$\left\{\begin{matrix} a + ar + ar^2 + \cdots\cdots \\ \frac{2}{3} + \frac{2}{3}\left[\left(\frac{2}{3}\right)^2\right]^1 + \frac{2}{3}\left[\left(\frac{2}{3}\right)^2\right]^2 + \cdots\cdots \end{matrix}\right\}$

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

$= \frac{a}{1-r}$

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

$= \frac{1}{2}\left[\frac{2/3}{1-4/9}\right]$

$= \boxed{\frac{3}{5}}$

TYPE 4

(9) From a sealed box containing a dozen apples, it was found that 3 apples are perished. Obtain the probability distribution of the no. of perished apples, when 2 apples are drawn at random. Also find the mean & variance.

$\rightarrow$ Let x be the no. of perished apples,

$x = 0, 1, 2, \ldots$

$p(x=0$, getting zero perished apples$) = \frac{{}^3C_0 \cdot {}^9C_2}{{}^{12}C_2} = \boxed{\frac{6}{11}}$

$P(x=1$, getting one perished apples$) = \frac{{}^3C_1 \cdot {}^9C_1}{{}^{12}C_2} = \boxed{\frac{9}{22}}$

$p(x=2$, getting two perished apples$) = \frac{{}^3C_2 \cdot {}^9C_0}{{}^{12}C_2} = \boxed{\frac{1}{22}}$

Distribution Table,

x	0	1	2
$p(x)$	$6/11$	$9/22$	$1/22$

We observe that, (i) $p(x) \geq 0$

(ii) $\Sigma p(x) = \frac{6}{11} + \frac{9}{22} + \frac{1}{22} \Rightarrow \boxed{\Sigma p(x) = 1}$

$\therefore p(x)$ is probability mass func.

Mean $\mu = \Sigma x \cdot p(x)$

$= 0 \cdot \frac{6}{11} + 1 \cdot \frac{9}{22} + 2 \cdot \frac{1}{22}$

$= \frac{9}{22} + \frac{2}{22}$

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

variance, $v = [\Sigma x^2 \cdot p(x)] - \mu^2$

$= \left[0^2\left(\frac{6}{11}\right) + 1^2\left(\frac{9}{22}\right) + 2^2\left(\frac{1}{22}\right)\right] - \left(\frac{1}{2}\right)^2$

$= \left[\frac{9}{22} + \frac{4}{22}\right] - \frac{1}{4}$

$= \frac{13}{22} - \frac{1}{4}$

$= \boxed{\frac{15}{44}}$

(10) A box contains 12 balls out of which 3 are white & 9 are red. A sample of 3 balls are selected at random from the box. Let x denote the no. of white balls from the sample. Find the distribution of x, also find mean, variance & sd.

→ Given:- x = no. of white balls

$x = 0, 1, 2, 3$

- $p(x=0$, getting no white ball$) = \frac{{}^3C_0 \cdot {}^9C_3}{{}^{12}C_3} = \frac{21}{55}$
- $p(x=1$, getting one white ball$) = \frac{{}^3C_1 \cdot {}^9C_2}{{}^{12}C_3} = \frac{27}{55}$
- $P(x=2$, getting two white ball$) = \frac{{}^3C_2 \cdot {}^9C_1}{{}^{12}C_3} = \frac{27}{220}$
- $P(x=3$, getting three white ball$) = \frac{{}^3C_3 \cdot {}^9C_0}{{}^{12}C_3} = \frac{1}{220}$

Distribution Table,

x	0	1	2	3
$P(x)$	$21/55$	$27/55$	$27/220$	$1/220$

we observed that,

(i) $p(x) \geq 0$

(ii) $\Sigma p(x) = \frac{21}{55} + \frac{27}{55} + \frac{27}{220} + \frac{1}{220}$

$\boxed{\Sigma p(x) = 1}$

∴ $p(x)$ is probability mass fn

- Mean $\mu = \Sigma x \cdot p(x)$

$= 0(21/55) + 1(27/55) + 2(27/220) + 3(1/220)$

$= \boxed{3/4}$

- variance $= \Sigma x^2 \cdot p(x)$

$= 1(27/55) + 4(27/220) + 9(1/220)$

$= \boxed{\frac{81}{176}}$

- sd $= \boxed{0.6784}$

TYPE - NOT IMPORTANT

(11) A random variable x take the values $-3, -2, -1, 0, 1, 2, 3$ such that $p(x=0) = p(x<0)$ & $P(x=-3) = P(x=-2) = p(x=-1) = p(x=2) = p(x=3)$. Find the probability distribution.

x	-3	-2	-1	0	1	2	3
$P(x)$	P_1	P_2	P_3	P_4	P_5	P_6	P_7

→ • $p(x=0) = p(x<0)$

$P_4 = P(-3) + P(-2) + P(-1)$

$P_4 = P_1 + P_2 + P_3 \longrightarrow (1)$

• Given, $p(x=-3) = p(x=-2) = p(x=-1) = p(x=2) = p(x=3)$

$P_1 = P_2 = P_3 = P_5 = P_6 = P_7 \rightarrow (2)$

• WKT,

$P_1 + P_2 + P_3 + P_4 + P_5 + P_6 + P_7 = 1 \rightarrow (3)$

Substitute (2) in (1)

$P_4 = 3P_1$

Substitute (2) in (3)

$9P_1 = 1$

$\boxed{P_1 = 1/9}$

$P_4 = 3P_1$

$\boxed{P_4 = 1/3}$

x	-3	-2	-1	0	1	2	3
$P(x)$	$1/9$	$1/9$	$1/9$	$3/9$	$1/9$	$1/9$	$1/9$

* BINOMIAL DISTRIBUTION

• BERNOULI'S TRIAL :- A random experiment with only 2 possible outcomes named as success & failure is called Bernouli trial.

• BERNOULI'S THEOREM :- "The probability of x-success in 'n' trials is given by $p(x) = {}^nC_x\, p^x q^{n-x}$, when

p - probability of success

q - " - failure

• Note:-

(i) p(atleast 2)

$p(x \geq 2) = p(2) + p(3) + p(4) + \dots + p(n)$

(OR)

WKT, $P(2) + P(3) + P(4) + \dots + p(n) = 1 - [p(0) + p(1)]$

(ii) p(atleast 3)

$p(x \geq 3) = p(3) + p(4) + p(5) + \dots + p(n)$

(OR)

WKT, $p(3) + p(4) + p(5) + \dots + p(n) = 1 - [p(0) + p(1) + p(2)]$

(iii) p(atmost 2)

$p(x \leq 2) = p(0) + p(1) + p(2)$

• NOTE :- • Mean (μ) = np

• variance (v) = npq

• std deviation (σ) = $\sqrt{v}$

TYPE I :-

① When a coin is tossed 4 times, find the probability of getting exactly one head, atmost three heads, atleast two heads.

→ Given, $n=4$

p (getting head) $= \frac{1}{2}$ → probability of sucess

$\cdot\ p+q=1$

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

• Let x denotes, no. of heads

WKT, probability of 'x' success in 'n' trials,

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

$p(x) = {}^{4}C_{x} \left[\frac{1}{2}\right]^{x} \left[\frac{1}{2}\right]^{4-x}$

$p(x) = {}^{4}C_{x} \cdot \frac{1}{2^{x+4-x}}$

$\boxed{p(x) = \frac{1}{16}\,{}^{4}C_{x}}$ → probability mass fn

(i) p(exactly one head) $= \frac{1}{16}\,{}^{4}C_{1}$

$= \frac{1}{16}(4)$

$= \boxed{\frac{1}{4}}$

(ii) p(atmost 3 heads) $= p(x \leqslant 3)$

$= p(0)+p(1)+p(2)+p(3)$

$= \frac{1}{16}\,{}^{4}C_{0} + \frac{1}{16}\,{}^{4}C_{1} + \frac{1}{16}\,{}^{4}C_{2} + \frac{1}{16}\,{}^{4}C_{3}$

$= \frac{1}{16}\left[{}^{4}C_{0} + {}^{4}C_{1} + {}^{4}C_{2} + {}^{4}C_{3}\right]$

$= \frac{1}{16}\left[1+4+6+4\right]$

$p(x \leqslant 3) = \boxed{\frac{15}{16}}$

(iii) p(atleast 2 heads) $= p(x \geqslant 2)$

$= p(2)+p(3)+p(4)$

$= \frac{1}{16}\left[{}^{4}C_{2} + {}^{4}C_{3} + {}^{4}C_{4}\right]$

$= \boxed{\frac{11}{16}}$

② The probability that a pen manufactured by a factory be defective is $\frac{1}{10}$. If 12 such pens are manufactured, what is the probability that

(i) exactly 2 are defective

(ii) atleast 2 are defective

(iii) none of them are defective

→ Let x denotes, no. of defective pens,

$p = \frac{1}{10} = 0.1$

$q = 1 - \frac{1}{10} = \frac{9}{10} = 0.9$

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

$= {}^{12}C_{x} (0.1)^{x} (0.9)^{12-x}$ → probability mass func

(i) $p(x=2) = {}^{12}C_2\,(0.1)^2\,(0.9)^{10}$

$= 66\,(0.1)^2(0.9)^{10}$

$= \boxed{0.2301}$

(ii) $p(x \geq 2) = 1 - [p(0) + p(1)]$

$= 1 - \left[{}^{12}C_0\,(0.1)^0(0.9)^{12} + {}^{12}C_1\,(0.1)^1(0.9)^{11}\right]$

$= \boxed{0.341}$

(iii) $p(x=0) = {}^{12}C_0\,(0.1)^0\,(0.9)^{12-0}$

$= \boxed{0.2824}$

---

(3) No. of telephone lines busy at an instant of time is a binomial variate with probability 0.1 that a line is busy. If 10 lines are chosen at random, Find,

(i) no line is busy
(ii) all lines are busy
(iii) atleast one line is busy
(iv) atmost two lines are busy

→ Let x denote no. of telephone lines busy.

$p = 0.1$

$q = 1 - 0.1 = 0.9$

$n = 10$

$p(x) = {}^{10}C_x\,(0.1)^x\,(0.9)^{10-x}$

P.M. func

(i) p(no line is busy)

$p(x=0) = {}^{10}C_0\,(0.1)^0\,(0.9)^{10-0}$

$= \boxed{(0.9)^{10}}$

(ii) p(all lines are busy)

$p(x=10) = {}^{10}C_{10}\,(0.1)^{10}\,(0.9)^0$

$= \boxed{(0.1)^{10}}$

(iii) $p(x \geq 1) = 1 - [p(0)]$

$= 1 - \left[{}^{10}C_0\,(0.1)^0\,(0.9)^{10}\right]$

$= 1 - [(0.9)^{10}]$

$= \boxed{0.6513}$

---

(4) A no. of telephone lines busy at an instant of time is a binomial variate, with prob 0.2 that a line is busy. If 10 lines are chosen at random. Find the probability that,

(i) 2 lines are busy
(ii) atmost 2 lines are busy
(iii) atleast 2 lines are busy

→ Let x denote no. of lines busy

$p = 0.2$

$q = 0.8$

$n = 10$

$p(x) = {}^{10}C_x\,(0.2)^x\,(0.8)^{10-x}$

PM func.

(i) $p(x=2) = {}^{10}C_2 (0.2)^2 (0.8)^8$

$= \boxed{0.30199}$

(ii) $p(x \leq 2) = p(0)+p(1)+p(2)$

$= {}^{10}C_0 (0.2)^0 (0.8)^{10} + {}^{10}C_1 (0.2)^1 (0.8)^9 + {}^{10}C_2 (0.2)^2 (0.8)^8$

$= \boxed{0.67779}$

(iii) $p(x \geq 2) = 1 - [p(0)+p(1)]$

$= 1 - [{}^{10}C_0 (0.2)^0 (0.8)^{10} + {}^{10}C_1 (0.2)^1 (0.8)^9]$

$= \boxed{0.6242}$

---

⑤ In a Quiz Contest of answering Yes or No, whats the probability of guessing atleast 6 answers correctly out of 10 questions asked. Also find the probability of the same, if there are 4 options for a correct answer.

→ There are only 2 options

- Let x denote no. of correct ans.

$p = \frac{1}{2}, q = \frac{1}{2}, n = 10$

- $p(x) = {}^nC_x p^x q^{n-x}$

$= {}^{10}C_x \left[\frac{1}{2}\right]^x \left[\frac{1}{2}\right]^{10-x}$

$= {}^{10}C_x \frac{1}{2^x} \frac{1}{2^{10-x}}$

$= \boxed{{}^{10}C_x \frac{1}{2^{10}}}$

p(guessing atleast 6 answers correctly)

$p(x \geq 6) = p(6)+p(7)+p(8)+p(9)+p(10)$

$= \frac{1}{2^{10}} [{}^{10}C_6 + {}^{10}C_7 + {}^{10}C_8 + {}^{10}C_9 + {}^{10}C_{10}]$

$= \frac{1}{2^{10}} [210+120+45+10+1] = \boxed{0.377}$

Case 2: there are 4 options

. Let x denote no. of correct answers

$p = \frac{1}{4}, q = \frac{3}{4}, n = 10$

$p(x) = {}^nC_x p^x q^{n-x}$

$= {}^{10}C_x \left[\frac{1}{4}\right]^x \left[\frac{3}{4}\right]^{10-x}$

$= \boxed{{}^{10}C_x \frac{3^{10-x}}{4^{10}}}$ P.M func

(i) p(guessing atleast 6 answers correctly)

$p(x \geq 6) = p(6)+p(7)+p(8)+p(9)+p(10)$

$= \frac{1}{4^{10}} [{}^{10}C_6 \cdot 3^4 + {}^{10}C_7 \cdot 3^3 + {}^{10}C_8 \cdot 3^2 + {}^{10}C_9 \cdot 3^1 + {}^{10}C_{10} \cdot 3^0]$

$= \boxed{0.01972}$

---

⑥ The probability that a person aged 60 years, will live upto 70 is 0.65, what is the probability that out of 10 persons aged 60 atleast 7 of them will live upto 70.

→ Let x denote the no. of person aged 60 will live upto 70

$p = 0.65, q = 0.35, n = 10$

- $p(x) = {}^{10}C_x (0.65)^x (0.35)^{10-x} \rightarrow$ PM func

$p(x \geq 7) = p(7)+p(8)+p(9)+p(10)$

$= {}^{10}C_7 (0.65)^7 (0.35)^3 + {}^{10}C_8 (0.65)^8 (0.35)^2 + {}^{10}C_9 (0.65)^9 (0.35)^1 + {}^{10}C_{10} (0.65)^{10} (0.35)^0$

$= \boxed{0.5139}$

TYPE 2

① In 800 families with 5 children each, how many families will be expected to have 3 boys? 5 girls? either 2 or 3 boys? atmost 2 girls? By assuming probability of boys & girls to be equal.

→ Let x denote the no. of boys

p = probability of getting boy $= \frac{1}{2}$

q = — " —— " — girl $= \frac{1}{2}$

$n = 5$ [i.e., for one family]

WKT, $p(x) = {}^5C_x \left[\frac{1}{2}\right]^x \left[\frac{1}{2}\right]^{10-x}$

$= \boxed{{}^5C_x \frac{1}{2^5}}$

we have 800 families,

$800\, p(x) = 800 \times \frac{1}{2^5} \cdot {}^5C_x$

$f(x) = 800 \times \frac{1}{32} \cdot {}^5C_x$

$\boxed{f(x) = 25 \cdot {}^5C_x}$

(i) 3 boys, $x = 3$

$f(3) = 25 \cdot {}^5C_3$

$= \boxed{250}$, families contains 3 boys

(ii) 5 girls, $x = 0$

$f(0) = 25 \cdot {}^5C_0$

$f(0) = \boxed{25}$

(iii) Either 2 (or) 3 boys : $x = 2$ or $x = 3$

$= f(2) + f(3)$

$= 25 \cdot {}^5C_2 + 25 \cdot {}^5C_3$

$= 25\left[{}^5C_2 + {}^5C_3\right]$

$= \boxed{500}$

(iv) atmost 2 girls

5 boys, 0 girl + 4 boys, 1 girl + 3 boys, 2 girls

$f(x=5) + f(x=4) + f(x=3)$

$= 25 \cdot {}^5C_5 + 25 \cdot {}^5C_4 + 25 \cdot {}^5C_3$

$= \boxed{400}$

② If the mean & standard deviation of the number of correctly answered questions in a test given to 4096 students are 2.5 & $\sqrt{1.875}$. Find an estimate of the no. of candidates answering correctly.

(i) 8 or more questions

(ii) 2 or less

(iii) 5 questions

→ Mean $\mu = 2.5$, std deviation $\sigma = \sqrt{1.875}$

$\boxed{np = 2.5}$

$\sqrt{npq} = \sqrt{1.875}$

$npq = 1.875$

$q = \frac{1.875}{2.5}$

$\boxed{q = 0.75}$

WKT, $p + q = 1$

$\boxed{p = 0.25}$

$n(0.25) = 2.5$

$n = \frac{2.5}{0.25}$

$\boxed{n = 10}$

WKT, probability of x sucess in 'n' trials

$$p(x) = {}^{n}C_x p^x q^{n-x}$$

$$p(x) = {}^{10}C_x (0.25)^x (0.75)^{10-x}$$

↳ P.M func of 1 student

for 4096 students,

$$4096\,p(x) = 4096\,{}^{10}C_x (0.25)^x (0.75)^{10-x}$$

$$\boxed{f(x) = 4096\,{}^{10}C_x (0.25)^x (0.75)^{10-x}}$$

or

$$f(x) = 4096\,{}^{10}C_x \cdot \frac{3^{10-x}}{4^{10}}$$

$$f(x) = \frac{1}{256}\,{}^{10}C_x \cdot 3^{10-x}$$

(i) 8 or more questions :-

$= f(8) + f(9) + f(10)$

$= \frac{1}{256}\left[{}^{10}C_8 \cdot 3^2 + {}^{10}C_9 \cdot 3^1 + {}^{10}C_{10} \cdot 3^0\right]$

$= \boxed{1.703 \approx 2 \text{ students}}$

(ii) 2 or less questions :-

$= f(0) + f(1) + f(2)$

$= \frac{1}{256}\left[{}^{10}C_0 \cdot 3^{10} + {}^{10}C_1\, 3^9 + {}^{10}C_2\, 3^8\right]$

$= \boxed{2152.8 \approx 2153 \text{ students}}$

(iii) 5 questions :-

$= f(5)$

$= \frac{1}{256}\left[{}^{10}C_5 \cdot 3^5\right] = \boxed{239 \text{ students}}$

TYPE 3 :-

(3) 4 coins are tossed 100 times & the following results were obtained. Fit a binomial distribution for the data & calculate the theoretical frequencies.

No. of heads	0	1	2	3	4
Frequency	5	29	36	25	5

→ Let x denote, no. of getting heads

• Mean $(\mu) = \frac{\Sigma f(x)}{\Sigma f} = \frac{0+29+72+75+20}{100} = \boxed{1.96}$

$np = 1.96$

$4p = 1.96$

$p = \frac{1.96}{4}$

$\boxed{p = 0.49}$

$p + q = 1$

$\boxed{q = 0.51}$

$$p(x) = {}^{4}C_x (0.49)^x (0.51)^{4-x}$$

$$\boxed{f(x) = 100\,{}^{4}C_x (0.49)^x (0.51)^{4-x}}$$

$f(0) = 100\,{}^{4}C_0 (0.49)^0 (0.51)^{4-0}$

$= 100(1)(1)(0.51)^4 = 6.765 \approx \boxed{7}$

$f(1) = 100\,{}^{4}C_1 (0.49)^1 (0.51)^3$

$= 25.999 \approx \boxed{26}$

- $f(2) = 100\, {}^4C_2 (0.49)^2 (0.51)^2$

  $= 37.47 \approx \boxed{37}$

- $f(3) = 100\, {}^4C_3 (0.49)^3 (0.51)^1$

  $= 24.0004 \approx \boxed{24}$

- $f(4) = 100\, {}^4C_4 (0.49)^4 (0.51)^0$

  $= 5.765 \approx \boxed{6}$

• POISSON DISTRIBUTION :-

→ When p is very small & n is very large, then we should use poisson distribution.

PM func $\boxed{p(x) = \frac{m^x e^{-m}}{x!}}$

Mean $(\mu) = m = np$

$(V) = m$

std deviation $(\sigma) = \sqrt{m}$

TYPE I

① 2% of fusers manufactured by a company are found to be defective, find the probability that 200 fuses contains (i) no defective fusers

(ii) 3 or more defective fusers

→ $n = 200$

Let x denote no. of defective fusers

$p = 2\%,\ 0.02$

$q = 0.98$

$m = np$

$= 200 \times 0.02$

$= \boxed{4}$

$\boxed{p(x) = \frac{4^x e^{-4}}{x!}}$ pmf

(i) $p(x=0) = \frac{4^0 e^{-4}}{0!}$

$= \boxed{0.0183}$

(ii) $p(x \geq 3) = 1 - [p(0) + p(1) + p(2)]$

$= 1 - [0.0183 + 0.073 + 0.146]$

$= \boxed{0.76}$

② If the probability of a bad reaction from a certain injection is 0.001, determine the chance that out of 2000 individuals, more than two will get a bad reaction.

→ Let x denote the no. of bad reaction

$p = 0.001$, $n = 2000$

$m = np$

$= \boxed{2}$

$p(x) = \frac{m^x . e^{-m}}{x!}$

$= \frac{2^x e^{-2}}{2!}$

(i) p(more than 2 will give bad reaction)

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

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

$= \boxed{0.3233}$

③ A communication channel receives independent pulses at the rate of 12 pulses per micro second. The probability of transmission error is 0.001 for each micro second. Compute the probabilities of

(i) no error during a micro second
(ii) one error per micro second
(iii) atleast one error per micro second
(iv) two errors
(v) atmost two errors.

→ Let x denote no. of transmission error,

$p = 0.001$, $n = 12$

$m = \boxed{0.012}$

$$p(x) = \frac{m^x e^{-m}}{x!}$$

$$p(x) = \frac{[0.012]^x e^{-0.012}}{x!}$$

(i) p(no error)

$$p(x=0) = \frac{(0.012)^0 \cdot e^{-0.012}}{0!}$$

$$= \boxed{0.988072}$$

(ii) $p(x=1) = e^{-0.012} \cdot \frac{(0.012)^1}{1!}$

$$= \boxed{0.01186}$$

(iii) $p(x \geq 1) = 1 - p(0)$

$= 1 - 0.988072$

$= \boxed{0.01193}$

(iv) $p(x=2) = \frac{(0.012)^2 \cdot e^{-0.012}}{2!}$

$$= \boxed{0.00072}$$

(v) $p(x \leq 2) = p(0) + p(1) + p(2)$

$$= e^{-0.012}\left[\frac{(0.012)^0}{0!} + \frac{(0.012)^1}{1!} + \frac{(0.012)^2}{2!}\right]$$

$= 0.9999914 \approx \boxed{1}$

④ A shop has 4 diesel generator sets which it hires every day. The demand for a gen set on an average is a poisson variate with value 5/2. Obtain the probability that on a particular day.

(i) there was no demand
(ii) a demand had to be refused.

→ Let x denote no. of demand

$m = 5/2 = \boxed{2.5}$

$$p(x) = \frac{e^{-2.5}(2.5)^x}{x!}$$

(i) p(no demand)

$$p(x=0) = e^{-2.5} \cdot \frac{(2.5)^0}{0!}$$

$$= \boxed{0.082}$$

(ii) $p(x > 4)$

$= 1 - [p(0) + p(1) + p(2) + p(3) + p(4)]$

$$= 1 - e^{-2.5}\left[\frac{(2.5)^0}{0!} + \frac{(2.5)^1}{1!} + \frac{(2.5)^2}{2!} + \frac{(2.5)^3}{3!} + \frac{(2.5)^4}{4!}\right]$$

$= \boxed{0.108822}$

⑤ The probability that a news reader commits no mistake in reading the news is $1/e^{3}$. Find the probability that on a particular news broadcast he commits.

(i) only 2 mistakes
(ii) more than 3 mistakes
(iii) atmost 3 mistakes

→ Let x denote no. of mistakes

given, $p(x=0) = \frac{1}{e^3}$

$$p(x) = \frac{m^x e^{-m}}{x!}$$

$$p(0) = e^{-m}$$

$$\boxed{m=3}$$

(i) $p(x=2) = \frac{3^2 e^{-3}}{2!}$

$$p(x=2) = \boxed{0.22404}$$

(ii) $p(x>3) = 1 - [p(0) + p(1) + p(2) + p(3)]$

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

$$= 1 - 0.6472$$

$$= \boxed{0.3527}$$

(iii) $p(x \le 3) = p(0) + p(1) + p(2) + p(3)$

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

$$= \boxed{0.6472}$$

TYPE 2

① The number of accidents in a year to taxi drivers in a city follows a Poison distribution with mean 3. Out of 1000 taxi drivers find approximately the number of the drivers with

(i) no accident in a year

(ii) more than 3 accidents in a year.

→ Let x denote the no. of accidents

$$\mu = m = \boxed{3}$$

$$p(x) = \frac{3^x e^{-3}}{x!}$$

There are 1000 taxi drivers

$$1000\,p(x) = 1000 \cdot \frac{3^x e^{-3}}{x!}$$

• p(no accidents) $= p(x=0)$

$$= 1000 \cdot \frac{3^0 e^{-3}}{0!}$$

$$= 49.7 \approx \boxed{50}$$

• p(more than 3 accidents) $= p(x>3)$

$$= 1 - [p(0) + p(1) + p(2) + p(3)]$$

$$f(x) = \cancel{1-}\,1000\left[1 - e^{-3}\left[\frac{3^0}{0!} + \frac{3^1}{1!} + \frac{3^2}{2!} + \frac{3^3}{3!}\right]\right]$$

$$= \boxed{352.7681}$$

② In a certain factory turning out razor blades, there is a small probability of 1/500 for any blade to be defective. The blades are supplied in packets of 10. Use poison distribution to calculate the approximate no. of packets containing,

① no defective

② one defective

③ two defective blades in a consignment of 10000 packets

→ Let x denote the no. of defective blades

$$p = \frac{1}{500} = 0.002, \quad n = 10$$

Since p is very small & n is very large we should use poison distribution

$$p(x) = \frac{m^x e^{-m}}{x!} \qquad m = np = 10(0.002) = 0.02$$

$$p(x) = \frac{(0.02)^x e^{-0.02}}{x!}$$

$$p(x) = 0.9802 \frac{(0.02)^x}{x!}$$

PMF of 1 packet

there are 10000 packets

$$f(x) = 10000\, p(x)$$

$$f(x) = 10000 \times 0.9802 \left(\frac{(0.02)}{2!}\right)^x$$

$$f(x) = 9802 \frac{(0.02)^x}{2!}$$

(i) No. of packet containing no defective $[x=0]$

$$f(0) = 9802 \frac{(0.02)^0}{0!} = 9802$$

(ii) Packets containing one defective blades $[x=1]$

$$f(1) = 9802 \frac{(0.02)}{1!} \approx 196$$

(iii) packets containing two defective blades $[x=2]$

$$f(2) = 9802 \frac{(0.02)^2}{2!} \approx 2$$

CONTINUOUS PROBABILITY DISTRIBUTION

If a random variable takes not countable finite no. of values.

Ex:- Counducting a survey on life of electric bulbs.

PROBABILITY DENSITY FUNCTION:-

If $f(x)$ is said to be probability density function, it satisfies the following condition :-

(i) $f(x) \geq 0$

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

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

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

TYPE I

① Find constant K such that $f(x) = \begin{cases} Kx^2 & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$ is a PDF. Also compute (i) $p(1 < x < 2)$

(ii) $p(x \leq 1)$

(iii) $p(x > 1)$

(iv) Mean

(v) Variance

→ $f(x)$ is pdf

$\therefore \int_0^3 Kx^2.dx = 1$

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

$K\left(\frac{27}{3}\right) = 1$

$\boxed{K = 1/9}$

$f(x) = \begin{cases} 1/9\, x^2 & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$

(i) $p(1 < x < 2) = \int_1^2 1/9\, x^2.dx = \frac{1}{9}\left[\frac{x^3}{3}\right]_1^2 = \frac{1}{27}[2^3 - 1^3] = \boxed{\frac{7}{27}}$

(ii) $p(x \leq 1) = \int_{-\infty}^{1} 1/9\, x^2.dx = \frac{1}{27}[1^3] = \boxed{\frac{1}{27}}$

(iii) $p(x > 1) = 1 - p(x \leq 1)$

$= 1 - \frac{1}{27}$

$= \boxed{\frac{26}{27}}$

(iv) $\int_{-\infty}^{3} x.f(x).dx = \int_0^3 x \frac{1}{9} x^2\, dx$

$= \frac{1}{9}\left[\frac{x^4}{4}\right]_0^3$

$= \frac{1}{36}[81 - 0] = \frac{81}{36} = \boxed{\frac{9}{4}}$

(v) variance $= \int_0^3 x^2.\left(\frac{1}{9}x^2\right).dx - \left(\frac{9}{4}\right)^2$

$= \frac{1}{9}\int_0^3 x^4.dx - \left(\frac{9}{4}\right)^2$

$= \frac{1}{9}\left[\frac{x^5}{5}\right]_0^3 - \frac{81}{36}$

$= \frac{1}{45}[243 - 0] - \frac{81}{16} = \boxed{\frac{27}{80}}$

(2) A random variable x has the following density function, $f(x) = \begin{cases} Kx^2 & -3<x<3 \\ 0 & \text{otherwise} \end{cases}$. Find k & also find

(i) $p(1 \leq x \leq 2)$

(ii) $p(x \leq 2)$

(iii) $p(x>1)$

$\rightarrow$ $\int_{-3}^{3} Kx^2 \, dx = 1$

$K\left[\frac{x^3}{3}\right]_{-3}^{3} = 1$

$\frac{K}{3}[3^3-(-3)^3] = 1$

$\frac{K}{3}[27+27] = 1$

$\frac{54}{3}K = 1$

$18K = 1$

$\boxed{K = 1/18}$

$f(x) = \begin{cases} \frac{1}{18}x^2 & -3<x<3 \\ 0 & \text{otherwise} \end{cases}$

(i) $p(1 \leq x \leq 2) = \int_{1}^{2} \frac{1}{18} x^2 . dx = \frac{1}{18(3)}[x^3]_1^2$

$= \frac{1}{54}[8-1] = \boxed{\frac{7}{54}}$

(ii) $p(x \leq 2) = \int_{-3}^{2} \frac{1}{18} x^2 . dx = \frac{1}{54}[x^3]_{-3}^{2} = \frac{1}{54}[8+27] = \boxed{\frac{35}{54}}$

(iii) $p(x>1) = \int_{1}^{3} \frac{1}{18} x^2 . dx = \frac{1}{54}[x^3]_1^3 = \frac{1}{54}[27-\cancel{1}] = \boxed{\frac{26}{54}}$

(3) A random variable x has density func. $f(x) = \frac{K}{1+x^2}$ $-\infty < x < \infty$. Determine K, hence evaluate (i) $p(x \geq 0)$

(ii) $p(0<x<1)$

$\rightarrow$ $f(x)$ is density func.

(i) $f(x) \geq 0$

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

$\int_{-\infty}^{\infty} \frac{K}{1+x^2} . dx = 1$

$K \int_{-\infty}^{\infty} \frac{1}{1+x^2} . dx = 1$

$K[\tan^{-1}x]_{-\infty}^{\infty} = 1$

$K[\pi/2 + \pi/2] = 1$

$\boxed{K = 1/\pi}$

(i) $p(x \geq 0) = \int_{0}^{\infty} f(x) \, dx$

$= \int_{0}^{\infty} \frac{1/\pi}{1+x^2} . dx$

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

$= \frac{1}{\pi}[\tan^{-1}\infty - \tan^{-1}0]$

$= \frac{1}{\pi}[\frac{\pi}{2} - 0]$

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

(ii) $p(0<x<1) = \int_{0}^{1} f(x) . dx$

$= \int_{0}^{1} \frac{1/\pi}{1+x^2} . dx$

$= \frac{1}{\pi}[\tan^{-1}1 - \tan^{-1}0]$

$= \frac{1}{\pi}[\frac{\pi}{4} - 0]$

$= \boxed{\frac{1}{4}}$

④ If the time 't' years required to complete a software project has pdf of the form $\begin{cases} kt(1-t) & 0 \le t \le 1 \\ 0 & \text{otherwise} \end{cases}$ find k, also probability that project will be completed in less than 4 months.

$\rightarrow$ $\int_0^1 kt(1-t).dt = 1$ $\qquad$ ~~$p(x<4) = \int_0$~~

$k\int_0^1 t - t^2.\ dt = 1$

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

$k\left[\frac{1}{2} - \frac{1}{3}\right] = 1$

$\boxed{k = 6}$

$\int_0^{0.33} 6t(1-t)dt$

$= 6\left[\frac{t^2}{2} - \frac{t^3}{3}\right]_0^{0.33}$

$= 6\left[\frac{0.33^2}{2} - \frac{0.33^3}{3}\right]$

$= 6\left[\frac{(1/9)}{2} - \frac{(1/27)}{3}\right]$

$= 6\left[\frac{1}{18} - \frac{1}{81}\right]$

$= 6\left[\frac{63}{1458}\right]$

$= \boxed{0.2592}$

⑤ Find k such that $f(x) = \begin{cases} kxe^{-x} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$ is a PDF. Find Mean.

• NOTE:-

In continuous probability distribution, we have

(i) exponential distribution

(ii) normal distribution

• Exponential Distribution:- The CPD having probability density function $f(x) = \begin{cases} \alpha e^{-\alpha x} & \text{for } x>0 \\ 0 & \text{otherwise, where } \alpha > 0 \end{cases}$

$$\text{Mean}(\mu) = \frac{1}{\alpha}$$

$$S.D = \frac{1}{\alpha}$$

TYPE I :

① If x is an exponential variate with $\mu=3$. Find $f(x)$

(i) $p(x>1)$

(ii) $p(x<3)$

→ $\mu = 3$

$1/\alpha = 3$

$\boxed{\alpha = 1/3}$

Function becomes $f(x) = \begin{cases} \frac{1}{3} e^{-\frac{1}{3}x} & \text{for } x>0 \\ 0 & \text{otherwise} \end{cases}$

(i) $p(x>1) = \int_1^{\infty} \frac{1}{3} e^{-\frac{1}{3}x} . dx$

$= \frac{1}{3} \left[\frac{e^{-x/3}}{-1/3} \right]_1^{\infty} . dx$

$= \frac{-1}{\infty} + \frac{1}{e^{1/3}}$

$= e^{-1/3}$

$= \boxed{0.7165}$

(ii) $p(x<3) = \frac{1}{3} \int_0^3 e^{-\frac{1}{3}x} . dx$

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

$= -\left[e^{-1} - 1 \right]$

$= 1 - e^{-1}$

$= \boxed{0.63212}$

② Find K such that $f(x) = \begin{cases} Kxe^{-x} & 0<x<1 \\ 0 & \text{otherwise} \end{cases}$ is a P.D.F. Find mean.

→ $\int_0^1 Kxe^{-x} . dx = 1$

$K \int_0^1 \underset{\text{I}}{x} . \underset{\text{II}}{e^{-x}} = 1$

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

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

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

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

$= K \left[\frac{-2}{e} + 1 \right] = 1$

$= K \left[\frac{-2+e}{e} \right] = 1$ $\boxed{K = \frac{e}{e-2}}$

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

$= \int_0^1 x \cdot \left[\frac{e}{e-2}\right] x e^{-x} \cdot dx$

$= \frac{e}{e-2} \int_0^1 \underset{\text{I}}{x^2} \cdot \underset{\text{II}}{e^{-x}} \cdot dx$

$= \frac{e}{e-2}\left[\frac{x^2 e^{-x}}{(-1)} - \frac{(2x)\, e^{-x}}{(-1)(-1)} + \frac{2 \cdot e^{-x}}{(-1)(-1)(-1)}\right]_0^1$

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

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

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

$= \frac{e}{e-2}\left[\frac{-1}{e} - \frac{2}{e} - \frac{2}{e} + \frac{2}{1}\right]$

$= \frac{e}{e-2}\left[\frac{-1-2-2+2e}{e}\right]$

$= \boxed{\frac{2e-5}{e-2}}$

③ If x is an exponential variate with mean 5. Find

(i) $p(0 < x < 1)$

(ii) $p(-\infty < x < 10)$

(iii) $p(x < 0$ or $x \geq 1)$

$\rightarrow$ Mean $= 5$

$\frac{1}{\alpha} = 5$

$\boxed{\alpha = \frac{1}{5}}$

given:- $f(x)$ is exponential distribution

(i) $p(0<x<1) = \int_0^1 \frac{1}{5} \cdot e^{1/5x} \cdot dx$

$= \frac{1}{5}\int_0^1 e^{-1/5x} \cdot dx$

$= \frac{1}{5}\left[\frac{e^{-1/5 x}}{-1/5}\right]_0^1$

$= -1\left[e^{-1/5(1)} - e^{-1/5(0)}\right]$

$= -\left[e^{-1/5} - e^0\right]$

$= 1 - e^{-1/5}$

$= 1 - 0.8187$

$= \boxed{0.18127}$

(ii) $p(-\infty < x < 10) = \int_{+0}^{10} \frac{1}{5} \cdot e^{-1/5 x} \cdot dx$

$= \frac{1}{5}\int_0^{10} e^{-1/5 x} \cdot dx$

$= -\frac{1}{\cancel{5}}\left[e^{-1/5(10)} - e^{-1/5(0)}\right]$

$= -\left[e^{-2} - 1\right]$

$= 1 - e^{-2}$

$= \boxed{0.86466}$

(iii) $p(x\leq 0 \text{ or } x\geq 1)$

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

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

$$= \int_{1}^{\infty} \frac{1}{5} e^{-1/5 x}.dx$$

$$= \frac{1}{5}\left[\frac{e^{-1/5 x}}{-1/5}\right]_{1}^{\infty}$$

$$= -1\left[e^{-1/5(\infty)} - e^{-1/5}\right]$$

$$= e^{-1/5}$$

$$= \boxed{0.8187}$$

④ The length of telephone conversation in a booth has been an exponential distribution & found on an average to be 5 minutes. Find the probability that a random call made from this booth (i) ends less than 5 minutes
(ii) between 5 & 10 minutes

→ given :- $f(x)$ is exponential distribution

Mean = 5

$\boxed{\alpha = 1/5}$

(i) $\int_{0}^{5} \frac{1}{5} e^{-1/5 x}$

$$= -1\left[e^{-1/5(5)} - e^{-1/5(0)}\right]$$

$$= 1 - e^{-1} = \boxed{0.6321}$$

(ii) $\int_{5}^{10} \frac{1}{5} e^{-1/5 x}$

$$= -1\left[e^{-1/5(10)} - e^{-1/5(5)}\right]$$

$$= -1\left[e^{-2} - e^{-1}\right]$$

$$= e^{-1} - e^{-2}$$

$$= \boxed{0.2325}$$

⑤ In a certain town the duration of a shower is exponentially distributed with mean 5 minutes. What is the probability that a shower will last for : (i) 10 mins or more
(ii) less than 10 mins
(iii) Between 10 & 12 mins.

$\mu = 5;\ \alpha = 1/5$

→ (i) $\int_{10}^{\infty} \frac{1}{5} e^{-1/5 x}$

$$= -1\left[e^{-1/5(\infty)} - e^{-1/5(10)}\right]$$

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

$$= +e^{2}$$

$$= \boxed{0.13}$$

(ii) $\int_{0}^{10} \frac{1}{5} e^{-1/5 x}$

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

$$= 1 - 0.13$$

$$= \boxed{0.87}$$

(iii) $\int_{10}^{12} \frac{1}{5} e^{-1/5 x}$

$$\rightarrow -1\left[e^{-1/5(12)} - e^{-1/5(10)}\right]$$

$$= e^{-2} - e^{-2.4}$$

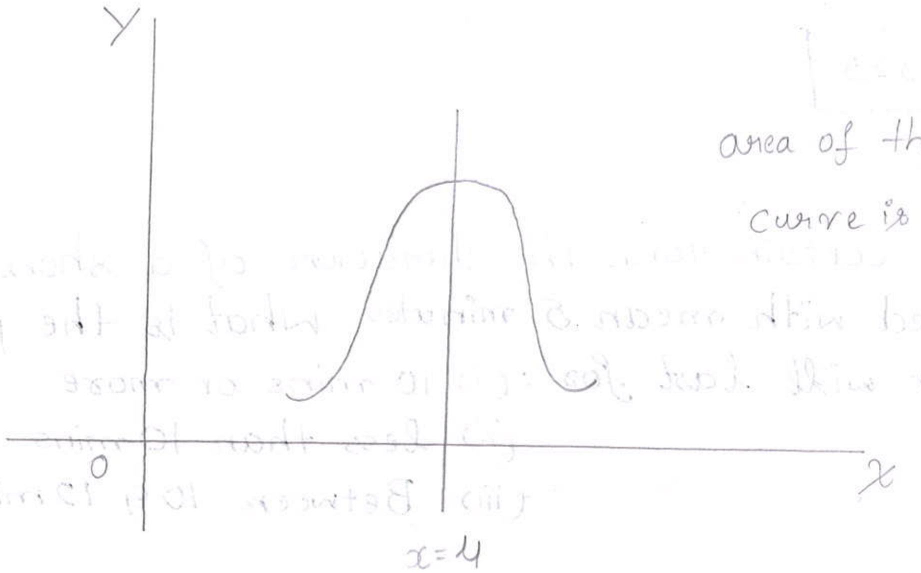
$$= \boxed{0.044}$$

*** Normal Distribution

The continuous probability distribution having pdf having

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

The function is symmetrical about $x = \mu$



area of this curve is 1

Consider, $p(a<x<b) = \int_a^b f(x).dx$

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

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

$z\sigma = x-\mu$

$x = z\sigma + \mu$

diff w.r.t z

(1) $\sigma + 0 = \frac{dx}{dz}$

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

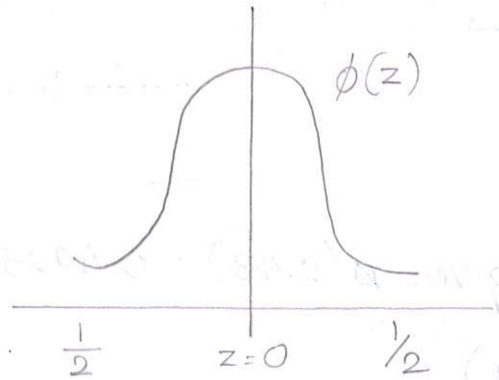
$z_1 = \frac{a-\mu}{\sigma}$

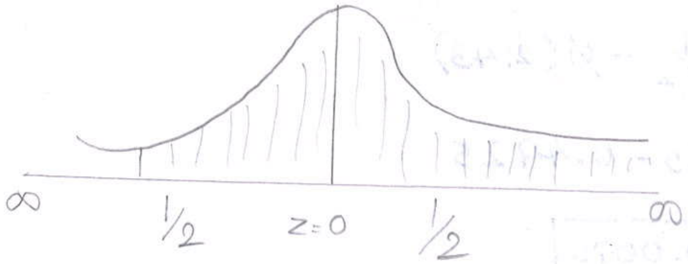
$z_2 = \frac{b-\mu}{\sigma}$

$$\int_{z_1}^{z_2} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}\, \sigma\, dz$$

$$\boxed{p(a<x<b) = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}}.\, dz}$$

Standard normal Pdf $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$





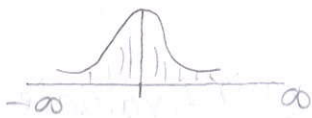
(i) $\int_{-\infty}^{\infty} \phi(z)\, dz = 1$

(ii) $\int_{-\infty}^{0} \phi(z).dz = 1/2$

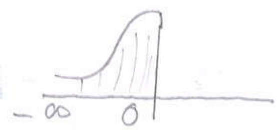
(iii) $\int_{0}^{\infty} \phi(z).dz = 1/2$

Equivalent forms

(i) $p(-\infty < z < \infty) = 1$



(ii) $p(-\infty < z < 0) = 1/2$

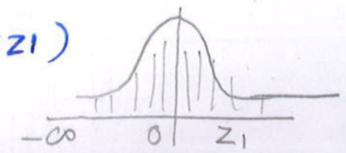


(iii) $p(0 < z < \infty) = 1/2$

(iv) $p(-\infty < z < z_1)$

$= p(-\infty < z < 0) + p(0 < z < z_1)$

$= \frac{1}{2} + \phi(z_1)$



TYPE 1 (z is given) :-

① Evaluate the following probabilities with the help of normal probability table.

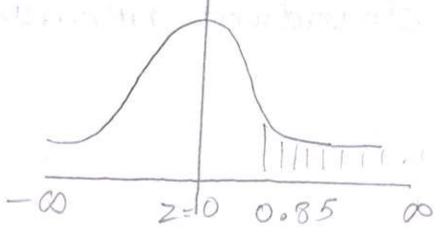
(i) Evaluate $p(z \geq 0.85)$ given $\phi(0.85) = 0.3023$

$\rightarrow$ $p(z \geq 0.85) = p(0 < z < \infty) - p(0 < z < 0.85)$

$= \frac{1}{2} - \phi(0.85)$

$= 0.5 - 0.3023$

$= \boxed{0.1977}$



(ii) Evaluate $p(z \leq -2.43)$ given $\phi(2.43) = 0.4925$

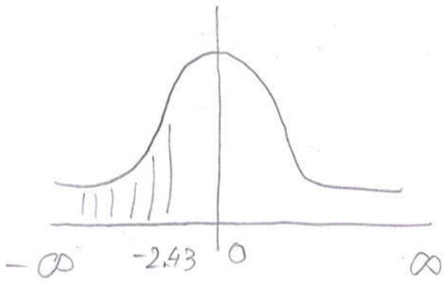
$\rightarrow$ $p(z \leq -2.43) = p(z \geq 2.43)$

$= p(0 < z < \infty) - p(0 < z < 2.43)$

$= \frac{1}{2} - \phi(2.43)$

$= 0.5 - 0.4925$

$= \boxed{0.0075}$

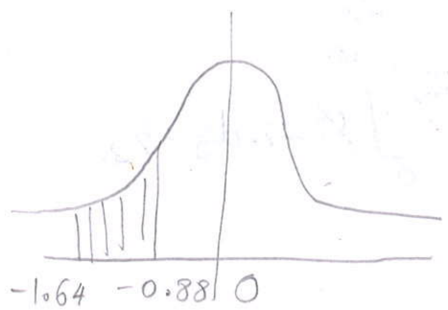


(iii) Evaluate $p(-1.64 \leq z \leq -0.88)$

$\rightarrow$ $= p(0 < z < 1.64) - p(0 < z < 0.88)$

$= 0.4495 - 0.3106$

$= \boxed{0.1389}$



(iv) $p(|z| \leq 1.94)$ given $p(1.94) = 0.4738$

$\rightarrow$ $p(|z| \leq 1.94) = p(-1.94 \leq z \leq 1.94)$

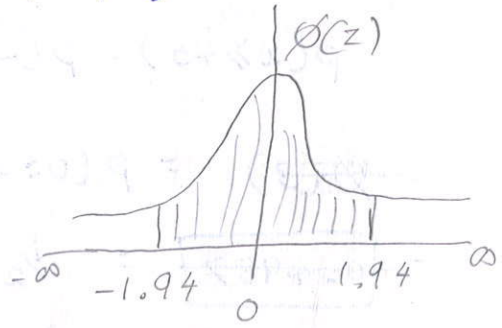
$= (-1.94 \leq z \leq 0) + (0 \leq z \leq 1.94)$

$= 2(0 \leq z \leq 1.94)$

$= 2\phi(1.94)$

$= 2(0.4738)$

$= \boxed{0.9476}$



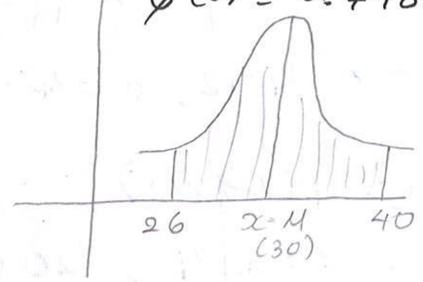
TYPE 2 [x is given]

① If x is a normal variate with mean 30 & sd 5. Find the probability that (i) $26 \leq x \leq 40$, given $\phi(0.8) = 0.2881$
(ii) $x \geq 45$

$\phi(2) = +0.4772$
$\phi(3) = 0.4987$

$\rightarrow$ $\mu = 30$, $\sigma = 5$

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



(i) $p(26 \leq x \leq 40)$

when $x = 26$, $z = \frac{x-\mu}{\sigma} = \frac{26-30}{5} = -0.8$

$x = 40$, $z = \frac{x-\mu}{\sigma} = \frac{40-30}{5} = 2$

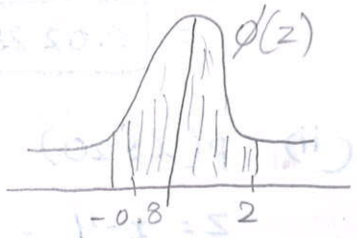
$p(26 \leq x \leq 40) = p(-0.8 \leq z \leq 2)$

$= (-0.8 \leq z \leq 0) + (0 \leq z \leq 2)$

$= p(0 \leq z \leq 0.8) + 0.4772$

$= 0.2881 + 0.4772$

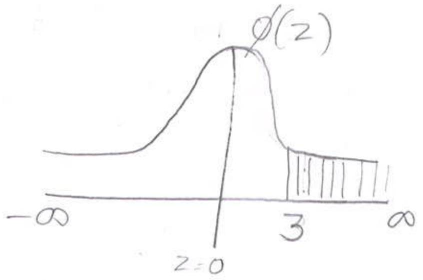
$= \boxed{0.7653}$



(ii) $p(x \geq 45)$

when $x=45$, $z = \frac{x-\mu}{\sigma}$, $= \frac{45-30}{5} = \frac{15}{5} = \cancel{2.53}$

$p(x \geq 45) = p(z \geq 3)$



$\cancel{=\phi(3)}$ $= p(0 \leq z \leq \infty) - p(0 \leq z \leq 3)$

$\cancel{= \boxed{0.4987}}$ $= \frac{1}{2} - \phi(3)$

$= \frac{1}{2} - 0.4987$

$= \boxed{0.0013}$

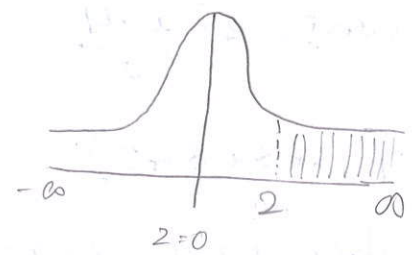
---

② If x is normally distributed with mean 12 & sd 4. Find the probability that (i) $(x \geq 20)$ (ii) $p(x \leq 20)$ $\phi(2) = 0.4772$

→ $\mu = 12$, $\sigma = 4$

(i) $p(x \geq 20)$

$z = \frac{x-\mu}{\sigma}$, $\frac{20-12}{4} = 2$



$p(x \geq 20) = p(z \geq 2)$

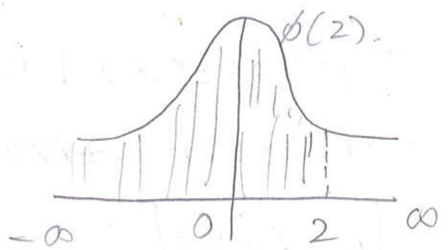
$= p(0 \leq z \leq \infty) - p(0 \leq z \leq 2)$

$= \frac{1}{2} - \phi(2)$

$= \boxed{0.0228}$

(ii) $p(x \leq 20)$

$z = \frac{x-\mu}{\sigma} = \frac{20-12}{4} = 2$



$= p(-\infty \leq z \leq 0) + \phi(2)$

$= \boxed{0.9772}$

***

TYPE 3

① The marks of 1000 students in an examination follows a normal distribution with mean 70 and sd 5. Find the no. of students whose marks will be

(i) less than 65 $\phi(1) = 0.3413$

(ii) more than 75

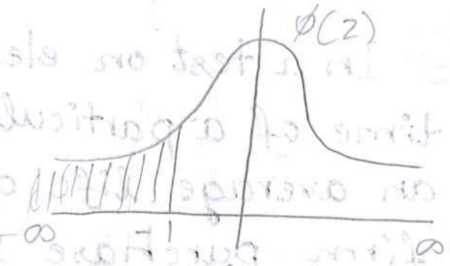
(iii) between 65 & 75

→ Let x denotes the marks of the students,

given, $\mu = 70$, $\sigma = 5$, WKT $z = \frac{x-\mu}{\sigma}$

(i) $p(x < 65)$

→ $x = 65$, $z = \frac{65-70}{5} = \boxed{-1}$



$p(x < 65) = p(z < -1)$

$= p(-\infty \leq z < 0) - p(0 < z < -1)$

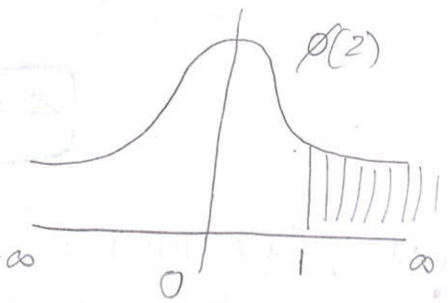
$= p(0 < z < \infty) - p(0 < z < 1)$

$= 0.5 - 0.3413$ for 1000 students

$= \boxed{0.1587} * 1000 = \boxed{159}$

(ii) $p(x > 75)$

→ $x = 75$, $z = \frac{75-70}{5} = \boxed{1}$



$p(x > 75) = p(z > 75)$

$= p(0 < z < \infty) - p(0 < z 1)$

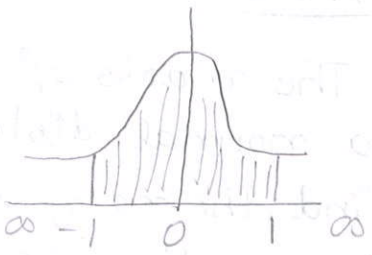
$= (0.5 - 0.3413) * 1000$

$= \boxed{159}$

(iii) $(65<x<75)$

→ when $x=65$, $\frac{65-70}{5} = -1$

$x=75$, $\frac{75-70}{5} = 1$



$p(\cancel{65}-1<z<1) = 2\phi(1)$

$= 2(0.3413)$

$= \boxed{0.6826} * 1000$

$= 682.6 \approx$ $\boxed{683 \text{ students}}$

② In a test on electric bulbs, it was found that the life time of a particular brand distributed normally with an average life of 2000 hours & SD 60 hours. If a firm purchase 2500 bulbs find the no. of bulbs that are likely to last for

(i) more than 2100 hours

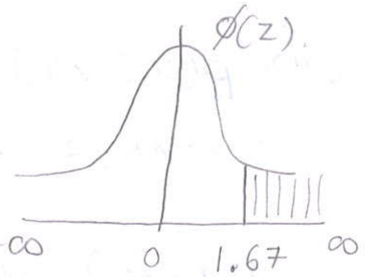
(ii) less than 1950 hours

(iii) between 1900 to 2100 hours

→ when $x=2100$, $Z = \frac{x-\mu}{\sigma}$

$Z = \frac{2100-2000}{60}$

$\boxed{Z = 5/3}$ or 1.67



(i) $p(x>2100) = p(z>1.67)$

$= 1/2 - \phi(1.67)$

$= 1/2 - 0.4525$

$= \boxed{0.0475} \times 2500 = \boxed{118.75}$

(ii) $p(x<1950)$

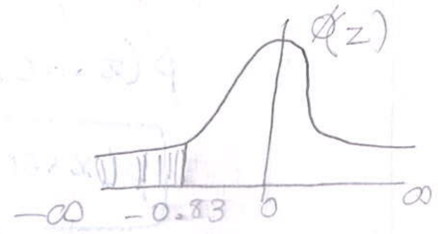
when $x=1950$, $= \frac{x-\mu}{\sigma} = \frac{1950-2000}{60} = \frac{-50}{60} = -0.833$

$p(x<1950) = p(z<0.833)$

$= 1/2 - \phi(0.833)$

$= \boxed{0.2033} * 2500$

$= \boxed{508.25}$



(iii) $p(1900<x<2100)$

when, $x=1900 = \frac{x-\mu}{\sigma} = \frac{100}{60}$

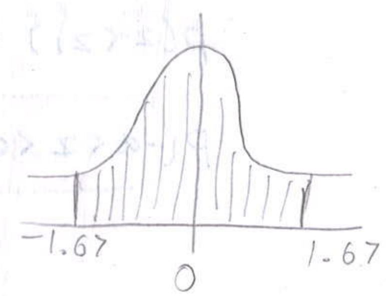
$= -1.67$

$= p(-1.67<z<1.67)$

$= 2\phi(1.67)$

$= 2(0.4525)$

$= \boxed{0.905} * 2500 = \boxed{2262.5}$



TYPE 4

① In an examination 7% of the students score less than 35% & 89% of the student score less than 60% marks. Find the Mean & SD if the marks are normally distributed. It is given that $p(z) = \frac{1}{\sqrt{2\pi}}\int_0^z e^{-\frac{z^2}{2}} . dz$

then $\phi(1.2263) = 0.39$

$\phi(1.4757) = 0.43$

→ Let x denotes the marks of the students

→ Let μ & σ be the mean & std deviation in normal distribution

given, 7% of the students sore less than 35% marks

$p(x<35) = 7\%$

$\boxed{p(x<35) = 0.07}$

* given, 89% of the students score less than 60% marks

$p(x<60) = 89\%$

$\boxed{p(x<60) = 0.89}$

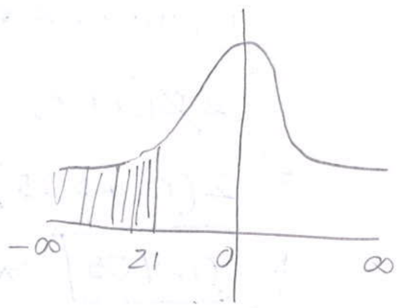
→ Consider $p(x<35) = 0.07$

if $x = 35$, $z = \frac{x-\mu}{\sigma} = \frac{35-\mu}{\sigma} = z_1$

$p(x<35) = 0.07$

$p(z<z_1) = 0.07$

$p(-\infty<z<0) - p(0<z<z_1) = 0.07$



$\phi(z_1) = 0.5 - 0.07$

$\boxed{\phi(z_1) = 0.43}$

→ Consider $p(x<60) = 0.89$

if $x = 60$, $z = \frac{x-\mu}{\sigma} = \frac{60-\mu}{\sigma} = z_2$

$p(x<60) = p(z<z_2) = 0.89$

$p(-\infty<z<0) + p(0<z<z_2) = 0.89$

$0.5 + \phi(z_2) = 0.89$

$\boxed{\phi(z_2) = 0.39}$

$\frac{35-\mu}{\sigma} = 0.43$; $\frac{60-\mu}{\sigma} = 0.39$

$35-\mu = 0.43(\sigma)$

$(-)\ 60-\mu = -0.39(\sigma)$

$-25 = -0.04\sigma$

$\sigma = \frac{25}{0.04}$

$\sigma = 625$

Given, $\phi(1.2263) = 0.39$

$\phi(1.2263) = \phi(z_2)$

$\boxed{z_2 = 1.2263}$

Given, $\phi(1.4757) = 0.43$

$\phi(1.4757) = \phi(z_1)$

$\boxed{z_1 = -1.475}$

$\frac{35-\mu}{\sigma} = -1.475$; $\frac{60-\mu}{\sigma} = 1.2263$

$35-\mu = (-1.475)\sigma$

$(-)\ 60-\mu = (1.2263)\sigma$

$-25 = -2.7013\sigma$

$\boxed{\sigma = 9.25}$

$60-\mu = 11.343$

$\boxed{\mu = 48.657}$

② In a normal distribution 31% of the items are under 45 & 8% of the items over 64. Find the mean & s.d.

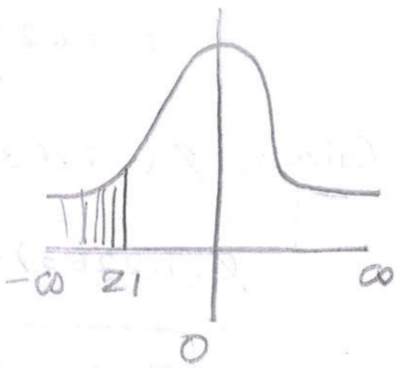
→ Let x denote the price of the items

given $p(x<45) = 0.31$

$p(x>64) = 0.08$

→ $p(x<45) = 0.31$

if $x=45$, $z = \frac{x-\mu}{\sigma} = \frac{45-\mu}{\sigma} = z_1$



$p(z<z_1) = 0.31$

$p(-\infty<z<0) - p(0<z<z_1) = 0.31$

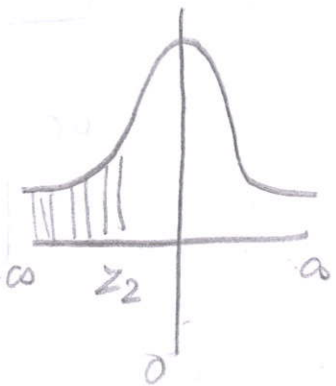
$\phi(z_1) = 0.5 - 0.31$

$\boxed{\phi(z_1) = 0.19}$

→ $p(x>64) = 0.08$

if $x=64$, $z = \frac{x-\mu}{\sigma} = \frac{64-\mu}{\sigma} = z_2$

$p(z<z_2) = 0.08$



~~$p(-\infty$~~ $p(-\infty<z<0) - p(0<z<z_2) = 0.08$

$\phi(z_2) = 0.5 - 0.08$

$\boxed{\phi(z_2) = 0.42}$

MODULE 3

SAMPLE THEORY

POPULATION:-

The large collection of individuals or numerical data is called population.
The no. of individuals in a population is called size, denoted by N.

SAMPLE:-

A pot selected from the population is called sample.
The no. of individuals in a sample is called sample size denoted by n.

CASE 1:- Random Sampling with replacement

Here the items are drawn one by one and put back to the population before the next draw

→ If N is the size of the population & n is the size of the sample then we have N^n samples.

Note :- (i) $\mu_x = \mu$ [Population Mean = Mean of sample mean]

(ii) $\sigma_{\bar{x}}^2 = \dfrac{\sigma^2}{n}$

or $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

CASE 2 :- Random sampling without replacement

Here the items are drawn one by one & put back to the population before the next draw

→ In this case there will be $^N C_n$ sample.

Note :- (i) $\mu_{\bar{x}} = \mu$ [population Mean = Mean of sample mean]

(ii) $\sigma_{\bar{x}} = \left[\dfrac{N-n}{N-1}\right]\dfrac{\sigma^2}{n}$

NUMERICALS :-

① A population consists of five numbers 2, 3, 6, 8, 11. Consider all possible samples of size 2 which can be drawn with replacement from this population. Find
(a) the mean & S.D of population
(b) the mean & SD of the sampling distribution of means.
(c) Considering samples without replacement, find the mean & S.D of the sampling distribution of means.

→ Given : population : 2, 3, 6, 8, 11

(a) • Mean $\mu = \dfrac{2+3+6+8+11}{5} = \dfrac{30}{5} = \boxed{6}$

• Variance $\sigma^2 = \dfrac{1}{N}\Sigma(x-\mu)^2$

$= \dfrac{1}{5}[(2-6)^2+(3-6)^2+(6-6)^2+(8-6)^2+(11-6)^2]$

$= \dfrac{1}{5}[16+9+0+4+25]$

$= \dfrac{54}{5}$

$\sigma^2 = 10.8$

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

(b) With replacement,

$N=5, \ n=2, \ N^n = 25$

(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)

Mean of each sample

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

x	2	2.5	3	4	4.5	5	5.5	6	6.5	7	8	8.5	9.5	11
f	1	2	1	2	2	2	2	1	2	4	1	2	2	1

· Mean of Sample mean &

$$\mu_{\bar{x}} = \frac{\Sigma f x}{\Sigma f}$$

$$= \frac{2+5+3+8+9+10+11+6+13+28+8+17+19+11}{25}$$

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

· variance, $\sigma_{\bar{x}}^2 = \frac{\Sigma f x^2}{\Sigma f} - (\mu_{\bar{x}})^2$

$$= \frac{2^2(1)+(2.5)^2 2+3^2(1)+4^2(2)+(4.5)^2(2)+(5)^2(2)+(5.5)^2(2)+6^2(1)+(6.5)^2(2)+7^2(4)+8^2(1)+(8.5)^2(2)+(9.5)^2(2)+(11)^2(1)}{25} - (6)^2$$

$$= \frac{1035}{25} - 36$$

$\sigma_{\bar{x}}^2 = 5.4 \rightarrow$ variance

$\boxed{\sigma_{\bar{x}} = \sqrt{5.4}} \longrightarrow$ std deviation

(c) Without replacement,

$N=5, n=2 \quad {}^N C_n = {}^5 C_2 = 10$

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

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

(6,8), (6,11)

(8,11)

Mean of each sample,

2.5, 4, 5, 6.5

4.5, 5.5, 7

7, 8.5

9.5

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

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

• variance,

$$\sigma_{\bar{x}}^2 = \left[\frac{N-n}{N-1}\right]\frac{\sigma^2}{n}$$

$$= \left[\frac{5-2}{5-1}\right]\frac{10.8}{2}$$

$$= \frac{3}{4} * \frac{10.8}{2}$$

$$\boxed{\sigma_{\bar{x}} = \sqrt{4.05}}$$

(2) Verify $\mu_{\bar{x}} = \mu$ with replacement $\mu_{\bar{x}} = \mu$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ $\qquad \sigma_{\bar{x}}^2 = \left[\frac{N-n}{N-1}\right]\frac{\sigma^2}{n}$ without replacement

[~~If~~ we get verify word & the question is for 8M, we need to follow Ist procedure]

→ II Method:-

• population

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

→ variance,

$$\sigma^2 = \frac{\Sigma(x-\mu)^2}{N}$$

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

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

(i) with replacement,

$N=5,\ n=2\ \ N^n = 5^2 = 25$

Mean of Sample Mean,

i.e., $\mu_{\bar{x}} = \mu$

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

Std deviation of sample mean,

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

$$= \frac{10.8}{2}$$

$$\boxed{\sigma_{\bar{x}} = \sqrt{5.4}}$$

(ii) without replacement,

$\mu = 5,\ n = 2,\ {}^{N}C_{n} = {}^{5}C_{2} = 10$

Mean of sample mean,

i.e., $\mu_{\bar{x}} = \mu$

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

Std deviation for sample mean,

$$\sigma_{\bar{x}}^2 = \left[\frac{N-n}{N-1}\right]\frac{\sigma^2}{n}$$

$$= \left[\frac{5-2}{5-1}\right]\frac{10.8}{2}$$

$$= 3/4 \times \frac{10.8}{2}$$

$$\boxed{\sigma_{\bar{x}} = \sqrt{4.05}}$$

② A population consists of ~~4~~ no. ~~of~~ 3, 7, 11, 15. Find the mean & S.D of the sampling distribution of means by considering samplings of size 2 with replacement. Find the μ & σ of the sampling distribution of the means without replacement & verify.

→ • population Mean $= \frac{3+7+11+15}{4} = \boxed{\mu = 9}$

• std of population $\sigma^2 = \frac{1}{N}\Sigma(x-\mu)^2$

$= \frac{1}{4}[(3-9)^2+(7-9)^2+(11-9)^2+(15-9)^2]$

$= \frac{36+4+4+36}{4}$

$\boxed{\sigma = \sqrt{20}}$

(i) with replacement,

$N=4$, $n=2$, $N^n = 4^2 = 16$ samples

(3,3) (3,7) (3,11) (3,15)
(7,3) (7,7) (7,11) (7,15)
(11,3) (11,7) (11,11) (11,15)
(15,3) (15,7) (15,11) (15,15)

Mean of samples,

3, 5, 7, 9
5, 7, 9, 11
7, 9, 11, 13
9, 11, 13, 15

x	3	5	7	9	11	13	15
f	1	2	3	4	3	2	1

Mean of Sample Mean,

$\mu_{\bar{x}} = \frac{\Sigma fx}{\Sigma f}$

$= \frac{3+10+21+36+33+26+15}{16}$

$= \frac{144}{16}$

$\boxed{\mu_{\bar{x}} = 9}$

Std dev of Mean of sample mean,

$\sigma_{\bar{x}}^2 = \frac{\Sigma x^2 f}{\Sigma f} - (\mu_{\bar{x}})^2$ (use this when frequency is there)

$= \frac{9(1)+25(2)+49(3)+36(4)+121(3)+169(2)+225(1)}{16} - (9)^2$

$= \frac{1456}{16} - 81$

$\sigma_{\bar{x}}^2 = 10$ → variance

$\boxed{\sigma_{\bar{x}} = \sqrt{10}}$ → std deviation

verification,

$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$

$= \frac{20}{2}$

$\boxed{\sigma_{\bar{x}} = \sqrt{10}}$

(ii) without replacement

$N=4,\ n=2,\ {}^{N}C_n = {}^{4}C_2 = 6 \rightarrow$ samples

(3,7), (3,11), (3,15)
(7,11), (7,15)
(11,15)

Mean of sample mean,

5, 7, 9
9, 11
13

$$\mu_{\bar{x}} = \frac{5+7+9+9+11+13}{6} = \boxed{9}$$

6 → no. of terms

· variance,

$$\sigma_{\bar{x}}^2 = \frac{\Sigma(x-\mu)^2}{{}^{N}C_n}$$ ↗ use this when there is no frequency

$$= \frac{(5-9)^2+(7-9)^2+(9-9)^2+(9-9)^2+(11-9)^2+(13-9)^2}{6}$$

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

$\sigma_{\bar{x}}^2 = 6.6 \rightarrow$ variance

$\boxed{\sigma_{\bar{x}} = \sqrt{6.6}} \rightarrow$ std deviation

verification,

$$\sigma_{\bar{x}}^2 = \left[\frac{N-n}{N-1}\right]\frac{\sigma^2}{n}$$

$$= \left[\frac{4-2}{4-1}\right]\frac{20}{2}$$

$$\boxed{\sigma_{\bar{x}} = \sqrt{6.6}}$$

③ population : 1, 2, 3. Consider all samples of size 2 which can be drawn with replacement from this population

(i) Find μ & σ of population

(ii) μ & σ of the sampling distribution of means

(iii) Considering samples without replacement

(iv) Considering samples with replacement

→ Given, population :- 1, 2, 3

Mean of population :- $\frac{1+2+3}{3}$

$\mu = 6/3 \quad \boxed{\mu = 2}$

Std deviation of Mean.

$$\sigma^2 = \frac{\Sigma(x-\mu)^2}{N}$$

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

$\sigma_{\bar{x}}^2 = 2/3$

$$\boxed{\sigma_{\bar{x}} = \sqrt{2/3}}$$

With replacement

$\mu = 3,\ n=2,\ N^n = 3^2 = 9$ samples

(1,1) (1,2) (1,3)
(2,1) (2,2) (2,3)
(3,1) (3,2) (3,3)

Mean of each sample

1, 1.5, 2
1.5, 2, 2.5
2, 2.5, 3

Frequency distribution :-

x	1	1.5	2	2.5	3
f	1	2	3	2	1

Mean of sample mean

$$\mu_{\bar{x}} = \frac{\Sigma f x}{f}$$

$$= \frac{1+3+6+5+3}{9}$$

$$\boxed{\mu_{\bar{x}} = 2}$$

Hence $\mu_{\bar{x}} = \mu$ verified

Std deviation of Mean of sample

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

$$= \frac{1(1) + 2.25(2) + 4(3) + 6.25(2) + 9(1)}{9} - (2)^2$$

$$\boxed{\sigma_{\bar{x}}^2 = 1/3} \qquad \boxed{\sigma_{\bar{x}} = \sqrt{1/3}}$$

$$\text{WKT, } \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{2/3}{2} = \frac{1}{3}$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \text{ verified}$$

Without replacement

$N=3$, $n=2$, ${}^{N}C_{n} = {}^{3}C_{2} = 3$ samples

(1,2) (1,3) (2,3)

Mean of each sample

1.5 , 2, 2.5

Mean of sample Mean $\mu_{\bar{x}} = \frac{1.5+2+2.5}{3} = \frac{6}{3} = 2$

$\therefore \mu_{\bar{x}} = \mu$ verified

Std deviation of Mean of samples

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

$$= \frac{(1.5-2)^2 + (2.2)^2 + (2.5-2)^2}{3}$$

$$\sigma_{\bar{x}}^2 = \frac{0.5}{3}$$

$$\boxed{\sigma_{\bar{x}} = \sqrt{1/6}}$$

$$\text{WKT } \sigma_{\bar{x}}^2 = \left[\frac{N-n}{N-1}\right] \frac{\sigma^2}{n}$$

$$= \left[\frac{3-2}{3-1}\right] \times \frac{2/3}{2}$$

$$\boxed{\sigma_{\bar{x}}^2 = 1/6}$$

$$\sigma_{\bar{x}}^2 = \left[\frac{N-n}{N-1}\right] \frac{\sigma^2}{n} \text{ verified}$$

④ The weights of 1500 ball bearings are normally distributed, with a mean of 635g & std dev 1.36g. If 300 random samples of size 36 are drawn from this population, determine expected mean & std dev of sampling distribution of means if sampling is done with & without replacement.

→ Given, $N = 1500$, $\mu = 635g$, $\sigma = 1.36g$, $n = 36$

(i) with replacement

Expected Mean $= \mu_{\bar{x}} = \mu$

$$\boxed{\mu_{\bar{x}} = 635}$$

Expected std dev $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$

$$= \frac{(1.36)^2}{36}$$

$$\boxed{\sigma_{\bar{x}} = 0.226}$$

(ii) without replacement

Expected Mean $\mu_{\bar{x}} = \mu$

$$\boxed{\mu_{\bar{x}} = 635}$$

Expected std dev,

$$\sigma_{\bar{x}}^2 = \left[\frac{N-n}{N-1}\right]\frac{\sigma^2}{n}$$

$$= \left[\frac{1500-36}{1500-1}\right]\frac{(1.36)^2}{36}$$

$$\boxed{\sigma_{\bar{x}} = 0.224}$$

⑤ The weights of 1500 ball bearings are normally distributed with a mean of 635g & std dev 1.36g. If 300 random samples of size 36 are drawn from this population, in the case of random sampling with replacement, find how many random samples would have their mean between (i) 634.76g & 635.249g

(ii) greater than 635.5g

(iii) less than 634.2g

(iv) less than 634.5g or more than 635.24g

→ Given, $N = 1500$, $\mu = 635g$, $\sigma = 1.36g$, $n = 36$

with replacement

(i) We have to find how many samples would samples would have their mean b/w 634.76g & 635.249g

$P(634.76g \leq \bar{x} \leq 635.249g)$ $\quad$ $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{(1.36)^2}{36} = 0.051$ $\quad$ $\boxed{\sigma_{\bar{x}} = 0.227}$

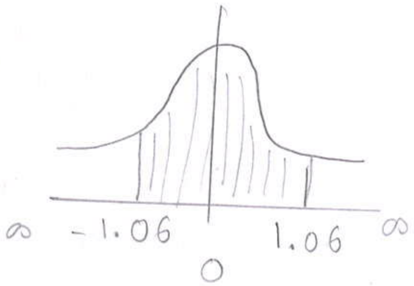
If $x = 634.76$, $Z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{634.76 - 635}{0.226} = \boxed{-1.06}$

If $x = 635.249$, $Z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{635.249 - 635}{0.226} = \boxed{1.06}$

$p(634.76 \leq \bar{x} \leq 635.24) = p(-1.06 \leq z \leq 1.06)$

$= p(-1.06 \leq z \leq 0)$

$+ p(0 \leq z \leq 1.06)$



$= 2\phi(1.06)$

$= 2(0.3554)$

$= \boxed{0.7108} * 300 \approx \boxed{213 \text{ samples}}$

(ii) we have to find how many samples would have their Mean greater than 635.5g

$p(\bar{x} > 635.5)$

If $\bar{x} = 635.5$, $z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{635.5 - 635}{0.227} = 2.203$

$p(\bar{x} > 635.5) = p(z > 2.203)$

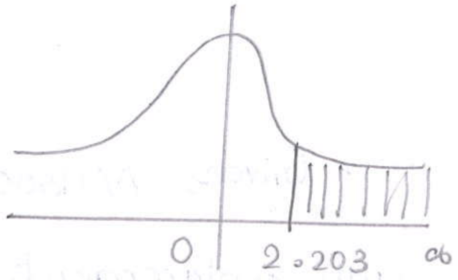
$= p(0 < z < \infty) - p(0 < z < 2.203)$

$= 0.5 - \phi(2.203)$

$= 0.5 - 0.4871$

$= \boxed{0.0129} * 300$

$= \approx \boxed{4 \text{ samples}}$



(iii) $p(\bar{x} < 634.2)$

If $\bar{x} = 634.2$, $z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{634.2 - 635}{0.227} = -3.52$

$p(\bar{x} < 634.2) = p(z < -3.52)$

$= p(-\infty < z < 0) - p(-3.54 < z < 0)$

$= \frac{1}{2} - \phi(3.52)$

$= 0.5 - 0.4998$

$= 0.0002$ probability of 1 sample

we have 300 samples

$= 300 \times 0.0002$

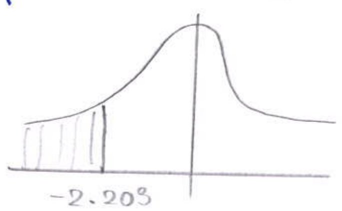
$= \boxed{0 \text{ samples}}$

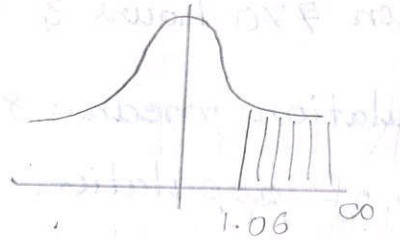
(iv) $p(\bar{x} < 634.5) + p(\bar{x} > 635.24)$

If $\bar{x} = 634.5$, $z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{634.5 - 635}{0.227} = -2.203$

If $\bar{x} = 635.24$, $z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{635.24 - 635}{0.227} = 1.06$

$p(z < -2.203) + p(z > 1.06)$





$\left(p(-\infty < z < 0) - p(-2.203 < z < 0)\right) + \left(p(0 < z < \infty) - p(0 < z < 1.06)\right)$

$\left(p(0 < z < \infty - p(0 < z < 2.203)\right) + \left(0.5 - \phi(1.06)\right)$

$= 0.5 - \phi(2.203) + (0.5 - \phi(1.06))$

$= 0.5 - 0.4871 + 0.5 - 0.3554$

$= 0.1575 * 300$

$\approx \boxed{47 \text{ samples}}$

⑥ Certain tubes manufactured by a company have mean life time of 800 hours & S.D of 60 hours. Find the probability that a random sample of 16 tubes taken from the group will have a mean life time

(a) between 790 hours & 810 hours

(b) less than 785 hours

(c) more than 820 hours

(d) between 770 hours & 830 hours.

→ population mean = 800 (μ)

S.D of population = 60 (σ)

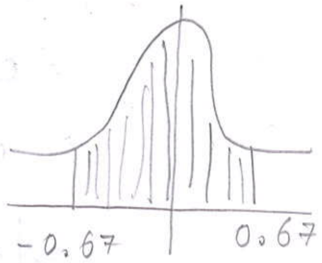
random sample = 16 (n)

(i) we have to find lifetime of 16 tubes b/w 790 & 810 hours

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{60}{4} = \boxed{15}$

If $\bar{x} = 790$, $z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} = \frac{790-800}{15} = -0.67$

$\bar{x} = 810$, $z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} = \frac{810-800}{15} = 0.67$



$p(790 < \bar{x} < 810) = p(-0.67 < z < 0.67)$

$= p(-0.67 < z < 0) + p(0 < z < 0.67)$

$= 2\phi(0.67)$

$= \boxed{0.4972} * 16 = \boxed{8}$

(ii) we have to find lifetime of 16 tubes less than 785 hours

$p(\bar{x} < 785)$

if $\bar{x} = 785$; $z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} = \frac{785-800}{15} = \frac{-15}{15} = 1$

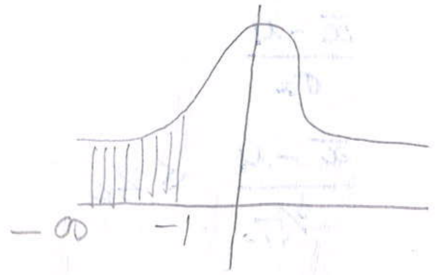
$p(\bar{x} < 785) = p(z < -1)$

$= p(-\infty < z < 0) + p(-1 < z < 0)$

$= p(0 < z < \infty) - p(0 < z < 1)$

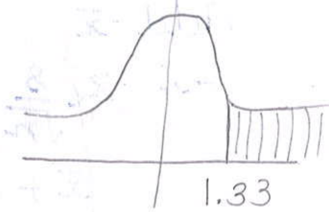
$= 0.5 - \phi(1)$

$= \boxed{0.1587} = \boxed{2.539}$



(iii) $p(\bar{x} > 820)$

If $\bar{x} = 820$, $z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} = \frac{820-800}{15} = \frac{20}{15} = 1.33$



$p(\bar{x} > 820) = p(z > 1.33)$

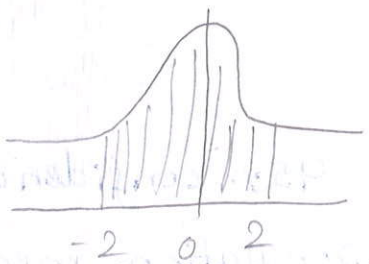
$= p(0 < z < \infty) - p(0 < z < 1.33)$

$= 0.5 - \phi(1.33)$

$= \boxed{0.0918} * 16 = \boxed{1.4688}$

(iv) If $\bar{x} = 770$, $z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} = \frac{770-800}{15} = -2$

$\bar{x} = 830$, $z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} = \frac{830-800}{15} = 2$



$= p(-2 < z < 2)$

$= p(-2 < z < 0) + p(0 < z < 2)$

$= 2\phi(2)$

$= \boxed{0.9544} = \boxed{15.27}$

TYPE - 3 [Confidence Interval (n = no. of samples)]

[$\bar{x}$, n , s , $\mu = ?$]

- $\bar{x}$ → mean of each sample
- n → no. of sample
- s → std deviation of sample
- $\mu = ?$ → population mean

$$\frac{\bar{x}-\mu}{\sigma_{\bar{x}}}$$

$$\frac{\bar{x}-\mu}{s/\sqrt{n}} \leq Z_\alpha$$

$$-Z_\alpha \leq \frac{\bar{x}-\mu}{s/\sqrt{n}} \leq Z_\alpha$$

$$-\frac{s}{\sqrt{n}} Z_\alpha \leq \bar{x}-\mu \leq Z_\alpha$$

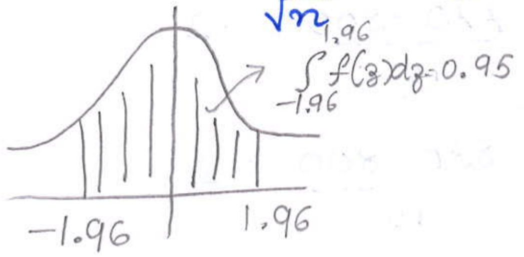
Add $-\bar{x}$

$$-\bar{x} - Z_\alpha \frac{s}{\sqrt{n}} \leq -\mu \leq -\bar{x} + \frac{s}{\sqrt{n}} Z_\alpha$$

$$\bar{x} + \frac{s}{\sqrt{n}} Z_\alpha \geq \mu \geq \bar{x} - \frac{s}{\sqrt{n}} Z_\alpha$$

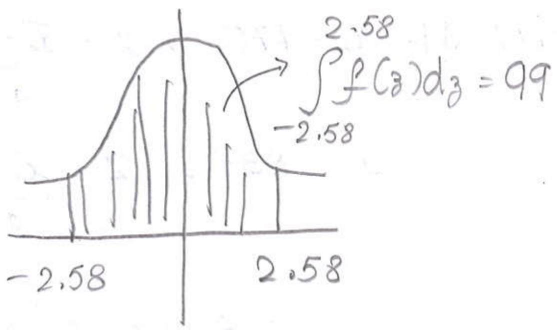
$$\boxed{\bar{x} - \frac{s}{\sqrt{n}} Z_\alpha \leq \mu \leq \bar{x} + \frac{s}{\sqrt{n}} Z_\alpha}$$

$$\bar{x} - \frac{s}{\sqrt{n}}(1.96) \leq \mu \leq \bar{x} + \frac{s}{\sqrt{n}}(1.96)$$



95% confidence that population mean will be similarly

$$\bar{x} - \frac{s}{\sqrt{n}}(2.56) \leq \mu \leq \bar{x} + \frac{s}{\sqrt{n}}(2.56)$$

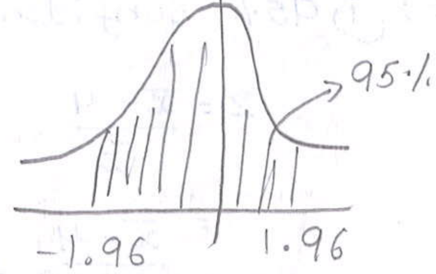


99% confidence that population mean will be in the internal

① A random sample of 400 terms choosen as infinite population is found to have mean of 82. If std deviation of 18. (i) Find 95% confidence limits. (ii) 99%.

→ $n = 400$, $\bar{x} = 82$, $s = 18$, $\mu = ?$

(i) $$\frac{\bar{x}-\mu}{\sigma_{\bar{x}}} \leq 1.96$$



$$\frac{\bar{x}-\mu}{s/\sqrt{n}} \leq 1.96$$

$$-1.96 \leq \frac{\bar{x}-\mu}{s/\sqrt{n}} \leq 1.96$$

$$-1.96 \frac{s}{\sqrt{n}} \leq \bar{x}-\mu \leq 1.96 \frac{s}{\sqrt{n}}$$

Add $-\bar{x}$

$$-\bar{x} - 1.96 \frac{s}{\sqrt{n}} \leq -\mu \leq -\bar{x} + 1.96 \frac{s}{\sqrt{n}}$$

$$\bar{x} + 1.96 \frac{s}{\sqrt{n}} \geq \mu \geq \bar{x} - 1.96 \cdot s/\sqrt{n}$$

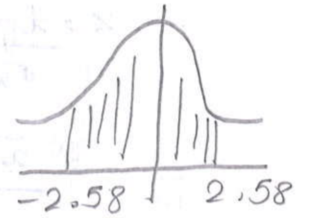
$$\bar{x} - 1.96 \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96\, s/\sqrt{n}$$

$$82 - 1.96 \frac{18}{\sqrt{400}} \leq \mu \leq 82 + 1.96 \cdot \frac{18}{\sqrt{400}}$$

$$\boxed{80.876 \leq \mu \leq 83.764}$$

95% confidence that population mean will be in interval.

(ii) 99% confidence limits



$$\bar{x} - 2.58 \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + 2.58 \frac{s}{\sqrt{n}}$$

$$82 - 2.58 \cdot \frac{18}{\sqrt{400}} \leq \mu \leq 82 + 2.58 \frac{18}{\sqrt{400}}$$

$$\boxed{79.696 \leq \mu \leq 84.304}$$

99% confidence that population will be in the interval

② The mean & S.D of marks scored by 100 students are 67.46 & 2.92. Find (i) 95% (ii) 99% interval for estimating mean marks of student interval/population.

→ (i) 95% confidence interval

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

$$= \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \leq 1.96$$

$$-1.96 \leq \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \leq 1.96$$

$$\bar{x} - 1.96 \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96 \frac{s}{\sqrt{n}}$$

$$67.45 - 1.96 \frac{2.92}{\sqrt{100}} \leq \mu \leq 67.45 + 1.96 \left(\frac{2.92}{\sqrt{100}}\right)$$

$$\boxed{66.88 \leq \mu \leq 68.02}$$

(ii) 99% confidence interval

$$z = \frac{\bar{x} - \mu}{\sigma_{x}}$$

$$= \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \leq 2.56$$

$$-2.56 \leq \frac{\bar{x} - \mu}{s/\sqrt{n}} \leq 2.56$$

$$\bar{x} - 2.56 \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + 2.56 \frac{s}{\sqrt{n}}$$

$$67.45 - 2.56 \left(\frac{2.92}{\sqrt{100}}\right) \leq \mu \leq 67.45 + 2.56 \left(\frac{2.92}{\sqrt{100}}\right)$$

$$\boxed{66.70 \leq \mu \leq 68.20}$$

③ The Mean & std dev of maximum loads supported by 60 cables are 11.09 tonnes & 0.73 tonnes respectively. Find 95%, 99% for mean of the maximum loads of all cables produced by the company.

→ $n = 60$, $\bar{x} = 11.09$, $s = 0.73$, $\mu = ?$

(i) 95% confidence limits :-

$$\bar{x} - 1.96 \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96 \frac{s}{\sqrt{n}}$$

$$11.09 - 1.96 \left(\frac{0.73}{\sqrt{60}}\right) \leq \mu \leq 11.09 + 1.96 \left(\frac{0.73}{\sqrt{60}}\right)$$

$$\boxed{10.91 \leq \mu \leq 11.27}$$

(ii) 99% confidence limits :-

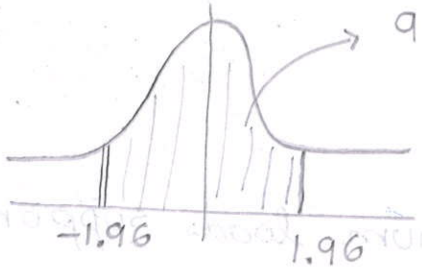
$$\bar{x} - 2.58 \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + 2.58 \frac{s}{\sqrt{n}}$$

$$11.09 - 2.58 \left(\frac{0.73}{\sqrt{60}}\right) \leq \mu \leq 11.09 + 2.58 \left(\frac{0.73}{\sqrt{60}}\right)$$

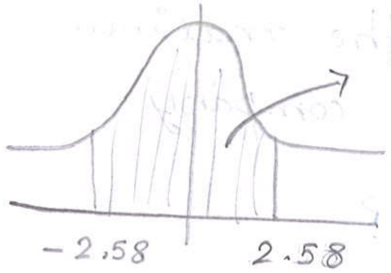
$$\boxed{10.85 \leq \mu \leq 11.33}$$

99% confidence that population Mean (μ) will be in this interval

** SIGNIFICANCE LEVEL :-



95% confidence interval
5% level of significance



99% confidence interval
1% level of significance

→ In order to take a decision regarding population, through a sample of population, we have to make certain assumptions called hypothesis

① The life of computers approximatey normally distributed with mean 800 hrs & SD 40 hrs. If random sample of 30 computers has an avg life of 788 hrs, test the hypothesis that the mean is equal to 800 hrs is acceptable at

(i) 5% level of significance

(ii) 1% level of significance

Also establish 95% confidence limits & 99% confidence limits.

→ given, $\mu = 800$ hrs

$\sigma = 40$ hrs

$n = 30$

$\bar{x} = 788$ hrs

Hypothesis :- population mean, $\mu = 800$ hrs

$$z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}}$$

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

$$= \frac{788-800}{\frac{40}{\sqrt{30}}}$$

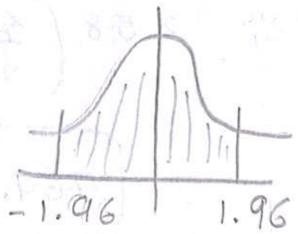
$$= -1.6432$$

$$\boxed{|z| = 1.6432}$$

Case 1: $|z| = 1.6432 < 1.96$

95% confidence that our hypothesis accepted

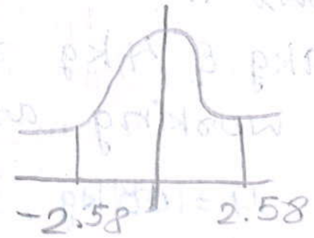
i.e., 5% level of significance



Case 2: $|z| = 1.6432 < 2.58$

99% confidence that our hypothesis accepted

i.e., 1% level of significance



(i) 95% confidence limits

$$\bar{x} - 1.96\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}$$

$$788 - 1.96\frac{40}{\sqrt{30}} \leq \mu \leq 788 + 1.96\frac{40}{\sqrt{30}}$$

$$\boxed{773.6861 \leq \mu \leq 802.3138}$$

(ii) 99% confidence limits

$$\bar{x} - 2.58\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 2.58\frac{\sigma}{\sqrt{n}}$$

$$788 - 2.58\left(\frac{40}{\sqrt{30}}\right) \leq \mu \leq 788 + 2.58\left(\frac{40}{\sqrt{30}}\right)$$

$$\boxed{769.158 \leq \mu \leq 806.814}$$

② A sugar factory expected sell sugar in 100kg bags. A sample of 144 bags taken from a day's output shows the average & S.D of weight of the bags are 99kg & 4kg respectively. Can we conclude that factory is working as per standards.

→ $\mu = 100kg$, $n = 144$, $\bar{x} = 99kg$, $s = 4kg$

Hypothesis, factory is working as per standards

i.e., $\mu = 100$

WKT, $Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$

$$= \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$= \frac{99 - 100}{4/\sqrt{144}}$$

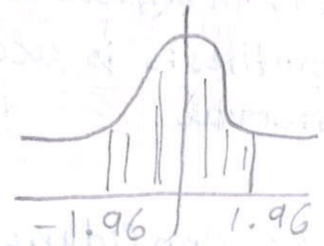
$$= \frac{-1}{4} \times 12$$

$$\boxed{|Z| = 3}$$

case(i):- $|Z| = 3 > 1.96$

95% confidence that our hypothesis is rejected

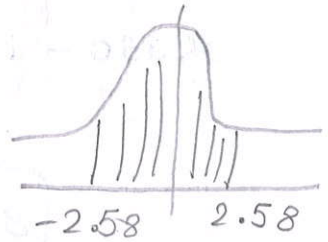
5% significance level



case(ii):- $|Z| = 3 > 2.58$

99% confidence that our hypothesis is rejected

1% significance level



③ A sample of 100 tyres is taken from a lot. The mean of life of tyres is found to be 39350km with SD 3260. Can it be considered as a true random sample from a population with mean life of 40000km. Use 0.05 level of significance. Establish 95% & 99% confidence limits.

→ given:- $n = 100$, $\bar{x} = 39350$ km, $s = 3260$

Hypothesis:- population mean, $\mu = 40000$ km

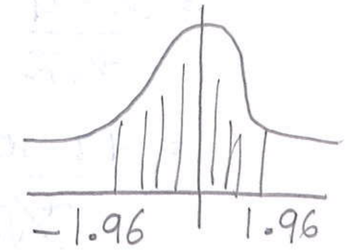
WKT $Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{39350 - 40,000}{3260/\sqrt{100}}$

$$\boxed{Z = -1.9938}$$

Case (i) $|Z| = 1.9938 > 1.96$

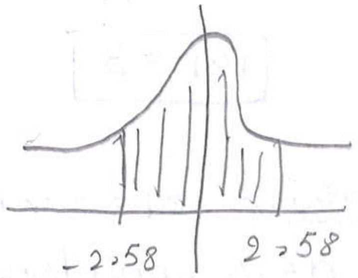
95% confidence that our hypothesis is rejected in this interval.

5% level of significance



Case (ii) $|Z| = 1.9938 < 2.58$

99% confidence that our hypothesis is accepted in this interval.



95% confidence limits

$$\bar{x} - 1.96\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96\frac{s}{\sqrt{n}}$$

$$39350 - 1.96 \times \frac{3260}{\sqrt{100}} \leq \mu \leq 39350 + 1.96 \times \frac{3260}{\sqrt{100}}$$

$$\boxed{38711.04 \leq \mu \leq 39988.96}$$

99% confidence limits

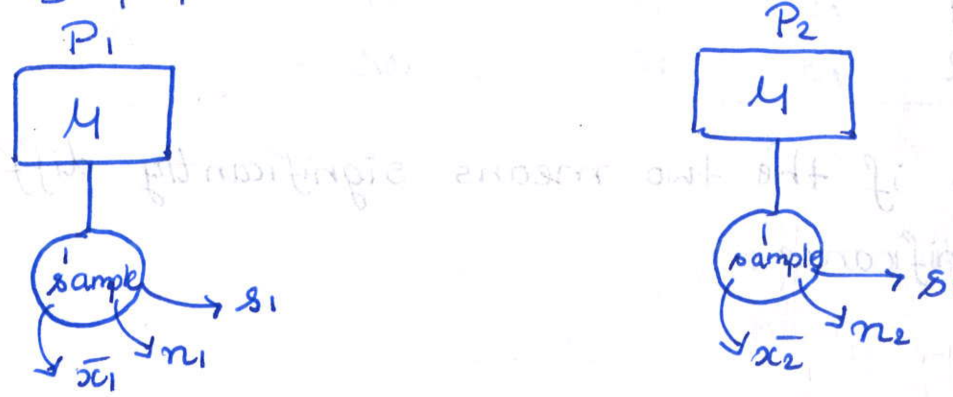
$$\bar{x} - 2.58\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + 2.58\frac{s}{\sqrt{n}}$$

$$39350 - 2.58 \times \frac{3260}{\sqrt{100}} \leq \mu \leq 39350 + 2.58 \times \frac{3260}{\sqrt{100}}$$

$$\boxed{38508.92 \leq \mu \leq 40191.08}$$

DIFFERENCE OF MEANS :-

When two different populations are given; we will take mean of population are equal.



To check the hypothesis,

$$\text{WKT}, \quad Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

$$= \frac{\bar{x} - \mu}{\sqrt{V(x)}}$$

$$= \frac{(\bar{x}_1 - \bar{x}_2) - (\mu - \mu)}{\sqrt{V(\bar{x}_1 - \bar{x}_2)}}$$

$$= \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{V(\bar{x}_1 - \bar{x}_2)}}$$

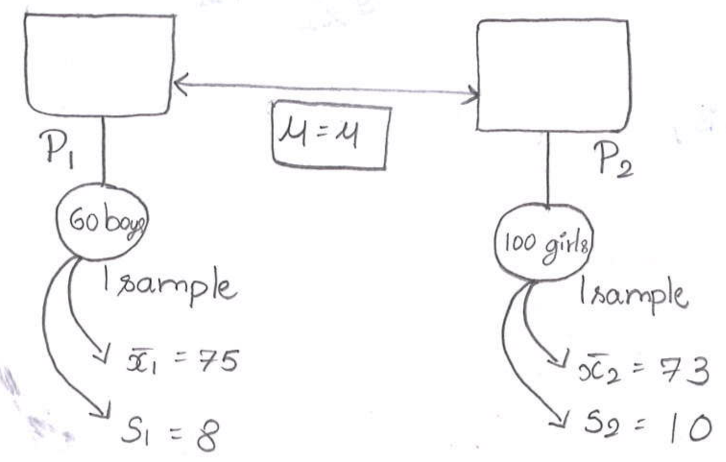
$$= \frac{(\bar{x}_1 - \bar{x}_2) - (\mu - \mu)}{\sqrt{\frac{s_1^2}{n_1}} + \sqrt{\frac{s_2^2}{n_2}}}$$

$$\boxed{Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}} + \sqrt{\frac{s_2^2}{n_2}}}}$$

① Intelligent Test were given to 2 groups of Boys & Girls.

	Mean	S.D	size of samples
Boy	75	8	60
Girl	73	10	100

Find out if the two means significantly differ at 5% significance.



→ Given, $n_1 = 60$, $\bar{x}_1 = 75$, $s_1 = 8$
$n_2 = 100$, $\bar{x}_2 = 73$, $s_2 = 10$

Hypothesis :- There is no significant difference b/w population mean

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$= \frac{75 - 73}{\sqrt{\frac{8^2}{60} + \frac{10^2}{100}}}$$

$$\boxed{|Z| = 1.39}$$

$\therefore$ $1.39 < 1.96$

our hypothesis is accepted 95% confidence that there is no difference in population Mean.

② In an elementary school examination, the mean grade of 32 boys was 72 with S.D of 8. while the mean grade of 36 girls was 75 with S.D of 6. Test the hypothesis that the performance of girls is better than boys.

→ Given $n_1 = 32$, $\bar{x}_1 = 72$, $S_1 = 8$
$n_2 = 36$, $\bar{x}_2 = 75$, $S_2 = 6$

Hypothesis :- There is no significant difference b/w population mean

(or)

there is no difference b/w performance of Boys & Girls

WKT,
$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

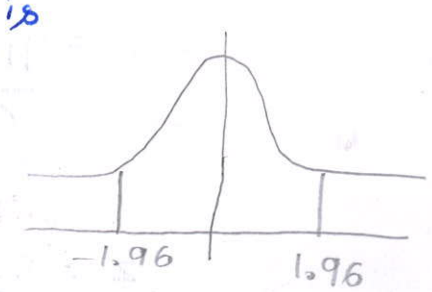
$$= \frac{72 - 75}{\sqrt{\frac{8^2}{32} + \frac{6^2}{36}}}$$

$$Z = -1.73$$

$$\boxed{|Z| = 1.73}$$

Case I :- $|Z| = 1.73 < 1.96$

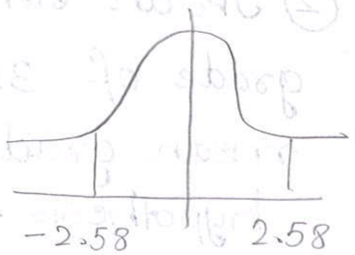
our hypothesis is accepted, there is no significant difference b/w the performance & boys & girls.



Case 2: $|Z| = 1.73 < 2.58$

our hypothesis is accepted

99% confidence that there is no significant difference b/w performance of boys & girls.



③ A random sample of 100 bulbs produced by a company 'A', showed a mean life of 1190 hrs, & S.D of 90 hours. Also sample of 75 bulbs produced by a company 'B', showed a mean life of 1230 hrs, & S.D of 120 hours. Is there a difference b/w mean life of bulbs produced by 2 companies at

(i) 5% level of significance

(ii) 1% level of significance

→ Given, $n_1 = 100$, $\bar{x}_1 = 1190$, $s.d = 90$

$n_2 = 75$, $\bar{x}_2 = 1230$, $s.d = 120$

Hypothesis :- There is no significant difference b/w mean life

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

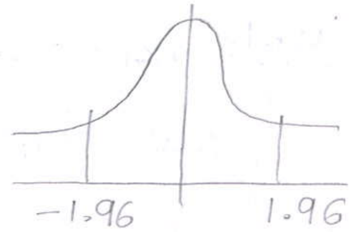
$$= \frac{1190 - 1230}{\sqrt{\frac{90^2}{100} + \frac{120^2}{75}}}$$

$$Z = -2.42$$

$$\boxed{|Z| = 2.42}$$

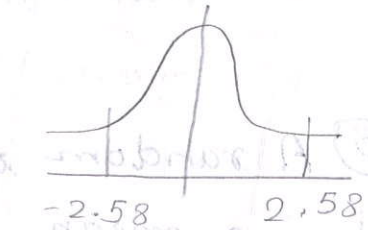
Case 1: $|Z| = 2.42 > 1.96$

our hypothesis is rejected in 95% confidence interval



Case 2 : $|Z| = 2.42 < 2.58$

our hypothesis is accepted in 99% confidence interval



④ The Mean of 2 large samples of 1000 & 2000 members are 168.75cm & 170cm respectively. Can the sample be regarded as drawn from the same population of S.D 6.25cm.

→ Hypothesis :- the sample are drawn from the same population of S.D 6.25

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

$$= \frac{168.75 - 170}{\sqrt{\frac{(6.25)^2}{1000} + \frac{(6.25)^2}{2000}}}$$

$$= \frac{-1.25}{\sqrt{(6.25)^2\left[\frac{1}{1000} + \frac{1}{2000}\right]}}$$

$$\boxed{|Z| = 5.16}$$

Case 1: $|z| = 5.16 > 1.96$

our hypothesis is rejected

Case 2- $|z| = 5.16 > 2.58$

our hypothesis is rejected

(5) A random sample of 1000 workers in a company has a mean wage of ₹50 per day & s.d of ₹15 per day. Another sample of 1500 workers has a mean wage ₹45 per day & S.D of ₹20 per day. Thus the mean rate of wages varies? Find 95% confidence limits for the difference of the mean wages.

→ Hypothesis:- Mean rate of wages doesn't varies

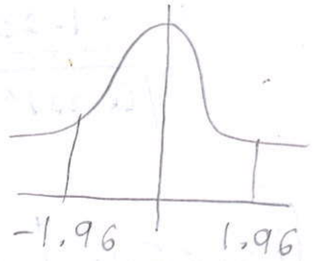
$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$= \frac{50-45}{\sqrt{\frac{15^2}{1000} + \frac{20^2}{1500}}}$$

$$\boxed{|z| = 7.13}$$

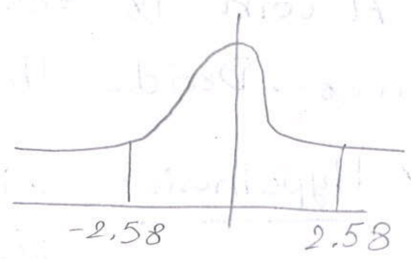
case 1: $z = 7.13 > 1.96$

our hypothesis is rejected in the 95% confidence interval



case 2: $z = 7.13 > 2.58$

our hypothesis is rejected in 99% confidence interval



Mean rate of wages varies

95% confidence limits

$$\left| \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \right| \leq 1.96$$

$$(\bar{x}_1 - \bar{x}_2) - 1.96\left(\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right) \leq \mu \leq (\bar{x}_1 - \bar{x}_2) + 1.96\left(\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)$$

$$= (50-45) - 1.96\left(\sqrt{\frac{15^2}{1000} + \frac{20^2}{1500}}\right) \leq \mu \leq (50-45) + 1.96\left(\sqrt{\frac{15^2}{1000} + \frac{20^2}{1500}}\right)$$

$$= \boxed{3.63 \leq \mu \leq 6.37}$$

* BINOMIAL DISTRIBUTION :-

Note:- Binomial distribution follows normal distribution

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

$$z = \frac{x - E(x)}{\sqrt{V(x)}}$$

$$z = \frac{x - np}{\sqrt{npq}}$$

① A coin is tossed 1000 times & head turns 540 times. Decide the hypothesis that the coin is unbiased.

→ Hypothesis :- coin is unbiased

i.e., $p = 1/2 \quad q = \frac{1}{2}$

Given, No. of samples = 1000

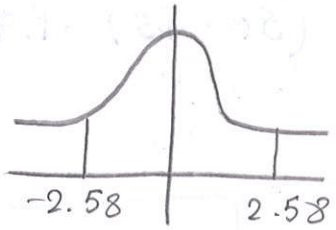
observed no. of success, $x = 540$

WKT, $Z = \dfrac{x - np}{\sqrt{npq}}$

$= \dfrac{540 - 1000(1/2)}{\sqrt{1000 \times (1/2)(1/2)}}$

$\boxed{Z = 2.53}$

$Z = 2.53 < 2.58$



② In 324 throws of 6 faced die, a odd no. turns up 181 times. Is it reasonable to think that die is unbiased one.

→ P = probability of getting odd no.s (1, 3, 5)

$P = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$

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

no. of samples $n = 324$

no. of success $x = 181$

$Z = \dfrac{x - np}{\sqrt{npq}}$

$= \dfrac{181 - 324(1/2)}{\sqrt{324(1/2)(1/2)}}$

$= \dfrac{19}{9}$

$\boxed{Z = 2.111 < 2.58}$

* PROPORTION :-

$Z = \dfrac{\bar{x} - \mu}{\sigma_{\bar{x}}}$

$= \dfrac{\bar{x} - E(\bar{x})}{\sqrt{V(\bar{x})}}$

$= \dfrac{\left(\frac{x}{n}\right) - E\left(\frac{x}{n}\right)}{\sqrt{V\left(\frac{x}{n}\right)}}$

$V\left(\frac{x}{n}\right) = \frac{1}{n^2} V(x)$

$= \frac{1}{n^2} npQ$

$= \frac{PQ}{n}$

$\boxed{Z = \dfrac{p - P}{\sqrt{\frac{PQ}{n}}}}$

① In a city consider sample of 500 people, 280 are tea drinkers & the rest are cofee drinkers. Can you assume that both coffee & tea are equally popular in the city at 5% level of significance

→ Hypothesis: coffee & tea are equally popular

$P = 1/2 \quad Q = \frac{1}{2} \quad n = 500$

proportion of 1 sample, $p = \left(\frac{x}{n}\right)$

$= \frac{280}{500}$

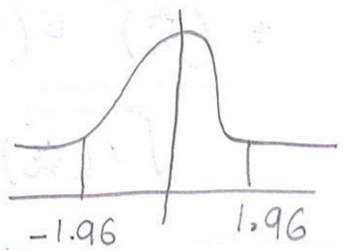
$\boxed{p = 0.56}$

WKT,

$$z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$$

$$= \frac{0.56 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{500}}}$$

$$\boxed{z = 2.68}$$



$z = 2.68 > 1.96$

our hypothesis is rejected

A random sample of 1000 engineering students from city Mysuru & 800 engineering from city Bengaluru were taken. It was found that 400 students in each of the sample were from payment Quota. Thus the data revealed a significance difference between 2 cities in respect of payment Quota students.

→

city 1

Mysuru

n_1 — 1000 stud.

1 sample

$x = 400$

$p_1 = \frac{x}{n_1} = \frac{400}{1000}$

$\boxed{p_1 = 0.4}$

city 2

Bengaluru

800 — n_2

1 sample

$x = 400$

$p_2 = \frac{x}{n_2} = \frac{400}{800}$

$\boxed{p_2 = 0.5}$

Hypothesis :- there is no significant difference b/w two cities

WKT, $Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

$= \frac{0.4 - 0.5}{\sqrt{(0.44)(0.56)\left[\frac{1}{1000} + \frac{1}{800}\right]}}$

$|Z| = 4.247 > 1.96$

> 2.58

our hypothesis is rejected

$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$

$= \frac{1000(0.4) + 800(0.5)}{1000 + 800}$

$\boxed{P = 0.44}$

$P + Q = 1$

$Q = 1 - P$

$\boxed{Q = 0.56}$

One type of aircraft is found to develop in 5 flights out of total of 100, and another type in 7 flights out of total of 200 flights. Is there a significant difference in 2 types of aircraft so far as engine defects are concerned.

→

aircraft 1

n_1 — 100 flights

1 sample

$p_1 = \frac{x_1}{n} = \frac{5}{100}$

$\boxed{p_1 = 0.05}$

aircraft 2

200 — n_2

1 sample

$p_2 = \frac{x_2}{n} = \frac{7}{200}$

$\boxed{p_2 = 0.035}$

Hypothesis :- there is no significant difference b/w the proportion of aircraft 1 & aircraft 2

$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

$= \frac{0.05 - 0.035}{\sqrt{(0.04)(0.96)\left[\frac{1}{100} + \frac{1}{200}\right]}}$

=

$P = \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2}$

$= \frac{0.05(100) + 0.035(200)}{100 + 200}$

$\boxed{P = 0.04}$

$P + Q = 1$

$Q = 1 - P$

$\boxed{Q = 0.96}$

MODULE 4

SAMPLING THEORY - 2

Student "t" test

(i) We should use student "t" test when

* sample size is $n \leq 30$
* population standard deviation σ is unknown

Formula :

WKT, $z = \dfrac{\bar{x} - \mu}{\sigma}$

$t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}} \Rightarrow t = \left(\dfrac{\bar{x} - \mu}{s}\right)\sqrt{n}$

$s = \dfrac{1}{n-1} \Sigma (x - \bar{x})^2$, $\bar{x} = \dfrac{1}{n} \Sigma x$

↳ degree of freedom

*TYPE I :-

① The individuals are chosen at random from population & heights in inches are found to be 63, 63, 66, 67, 68, 69, 70, 70, 71, 71, test hypothesis that the mean height of universe is 66 inches ($t_{0.05} = 2.262$ for 9 df)

→ Given, $n = 10$

Hypothesis ; Mean height of population

$\mu = 66$ inches

WKT, $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}} \rightarrow (*)$

Consider, $\bar{x} = \dfrac{1}{n} \Sigma x$

$= \frac{1}{10}[63+63+66+67+68+69+70+70+71+71]$

$\boxed{\bar{x} = 67.8}$

Consider, $s^2 = \frac{1}{n-1} \Sigma (x-\bar{x})^2$

$= \frac{1}{9}[(63-67.8)^2 + (63-67.8)^2 + \ldots \ldots \ldots]$

$= 9.067$

$\boxed{S = 3.011}$

(*) becomes : $t = \left(\frac{67.8-66}{3.011}\right)\sqrt{10} \Rightarrow 1.89$

$|t| = 1.89 < 2.262$

Our hypothesis is accepted.

(2) A certain stimulus administered to each of the 12 patients resulted in the following change in blood pressure 5, 2, 8, -1, 3, 0, 6, -2, 1, 5, 0, 4. Can it be concluded that stimulus will increase the blood pressure [$t_{0.05} = 2.201$ for 11 d.f]

$\rightarrow$ Given : $n = 12$

Hypothesis :- Stimulus will not increase the blood pressure, i.e, $\mu = 0$

WKT, $t = \frac{\bar{x}-\mu}{s/\sqrt{n}} \rightarrow$ (*)

Consider, $\bar{x} = \frac{1}{n}\Sigma x \Rightarrow \frac{5+2+8-1+3+0+6-2+1+5+0+4}{12}$

$\bar{x} = \frac{31}{12}$

$\boxed{\bar{x} = 2.583}$

Consider, $s^2 = \frac{1}{(n-1)} \Sigma (x-\bar{x})^2$

$= \frac{1}{12-1}[(5-2.583)^2 + (2-2.583)^2 + \ldots]$

$= \frac{1}{11}(104.918)$

$s^2 = 9.583$

$\boxed{S = 3.088}$

(*) becomes : $t = \frac{\bar{x}-\mu}{s/\sqrt{n}} \Rightarrow \frac{2.583-0}{3.088/\sqrt{2}}$

$\boxed{t = 2.89 > 2.801}$

Our hypothesis is rejected

$\therefore$ Stimulus will increase the blood pressure.

③ The height of 10 males of a given locality are found to be 70, 67, 62, 68, 61, 68, 70, 64, 64, 66 inches. Is it reasonable to believe to average height greater than 64 inches. Test the 5 % significance level.
[$t_{0.05}$ = 1.83 for 9 degree of freedom]

→ given: $n = 10$

Hypothesis: The mean height of the population is equal to 64 inches i.e., $\mu = 64$

WKT, $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ → ⊛

Consider, $\bar{x} = \dfrac{1}{n}\Sigma x = \dfrac{1}{10}[70+67+62+68+61+68+70+64+64+66]$

$= \dfrac{660}{10}$

$\boxed{\bar{x} = 66}$

Consider, $s^2 = \dfrac{1}{n-1}\Sigma(x-\bar{x})^2$

$= \dfrac{1}{66-1}[(70-66)^2 + (67-66)^2 + \dots (66-66)^2]$

$= \dfrac{1}{65}[16+1+16+4+25+4+0]$

$= \dfrac{90}{9}$

$s^2 = 10$

$\boxed{s = 3.16227}$

⊛ becomes, $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}} = \dfrac{66-64}{3.16227/\sqrt{10}} = \boxed{2}$

$\boxed{t = 2 > 1.83}$

since calculated t is greater than tabulated [$t_{0.05} = 1.83$] our hypothesis is rejected. ∴ The mean height of population mean is not equal to 64 inches.

④ A random sample of 10 boys had the following IQ's 70, 120, 110, 101, 88, 83, 95, 98, 107, 100. To this data, support the assumption of a population mean IQ of 100? Find the reasonable range in the which most of the mean IQ values of sample of 10 boys lie.

→ Given, $n = 10$

Hypothesis: population mean IQ is 100 i.e., $\mu = 100$

WKT, $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ → ⊛

Consider, $\bar{x} = \dfrac{1}{n}\Sigma x = \dfrac{1}{10}[70+120+110+101+88+83+85+98+107+100]$

$= \dfrac{972}{100}$

$= \boxed{97.2}$

Consider, $s^2 = \dfrac{1}{n-1}\Sigma(x-\bar{x})^2$

$= \dfrac{1}{10-1}[(70-97.2)^2 + (120-97.2)^2 + \dots + (100-97.2)^2]$

$s^2 = \dfrac{1}{9}[1833.6]$

$s^2 = 203.733$

$\boxed{s = 14.273}$

⊛ becomes, $= \dfrac{97.2-100}{14.273/\sqrt{10}} = 0.62$

$\boxed{t = 0.62 < 2.262}$

Our hypothesis is accepted

95% confidence,

$$|t| = \left|\frac{\bar{x}-\mu}{s/\sqrt{n}}\right| \leq 2.262$$

$$-2.262 \leq \frac{\bar{x}-\mu}{s/\sqrt{n}} \leq 2.262$$

$$-2.262\frac{s}{\sqrt{n}} \leq \bar{x}-\mu \leq 2.262\frac{s}{\sqrt{n}}$$

add $-\bar{x}$,

$$-\bar{x}-2.262\frac{s}{\sqrt{n}} \leq -\mu \leq -\bar{x}+2.262\frac{s}{\sqrt{n}}$$

$(* -1)$,

$$\bar{x}+2.262\frac{s}{\sqrt{n}} \geq \mu \geq \bar{x}-2.262\frac{s}{\sqrt{n}}$$

$$\boxed{\bar{x}-2.262\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x}+2.262\frac{s}{\sqrt{n}}}$$

or

$$\mu = x \pm 2.262\left(\frac{s}{\sqrt{n}}\right)$$

$$\mu = 97.2 \pm 2.262\left(\frac{14.273}{\sqrt{10}}\right)$$

$$97.2 - 2.262\left(\frac{14.273}{\sqrt{10}}\right) \ \& \ 97.2 + 2.262\left(\frac{14.273}{\sqrt{10}}\right)$$

$$\boxed{86.99 \ \& \ 107.41}$$

TYPE II :

① A random sample of size 16 has 53 as mean. The sum of squares of the deviation from the mean is 135. Can the sample be regarded as taken from population having 56 as mean? Obtain 95% confidence limit of the mean population.

$[t_{0.05} : 2.131 \text{ for } 15 \text{ df}]$

→ $n = 16$, $\bar{x} = 53$, $\Sigma(x-\bar{x})^2 = 135$

population $\rightarrow \mu = 56$

$16 \rightarrow n$, $\rightarrow \bar{x} = 53$

Hypothesis : population mean is 56

i.e., $\mu = 56$

WKT, $t = \dfrac{\bar{x}-\mu}{s/\sqrt{n}} \rightarrow (*)$

Consider, $s^2 = \dfrac{1}{n-1}\left[\Sigma(x-\bar{x})^2\right]$

$$= \frac{1}{16-1}[135]$$

$$= \frac{135}{15} \quad \Rightarrow s^2 = 9$$

$$\boxed{s = 3}$$

$(*)$ becomes,

$$t = \frac{53-56}{3/\sqrt{16}}$$

$$= \frac{-3}{3/4}$$

$$\boxed{|t| = 4}$$

$$\boxed{t = 4 > 2.131}$$

Our hypothesis is rejected.

95% confidence limits,

$$|t| = \left|\frac{\bar{x}-\mu}{s/\sqrt{n}}\right| \leq 2.131$$

$$\mu = \bar{x} \pm 2.131\left(s/\sqrt{n}\right)$$

$$= 53 \pm 2.131\left(3/\sqrt{16}\right)$$

$$= 53 \pm 2.131\left(3/4\right)$$

$$\mu = 53 - 2.131\left(3/4\right) \ \& \ 53 + 2.131\left(3/4\right)$$

$$51.40 \ \& \ 54.59$$

TYPE - I in t test :-

① A Group of boys & girls were given an intelligence test, the mean scores; S.D & no. of each group are :-

	BOYS	GIRLS
Mean	74	70
S.D	8	10
n	12	10

Is the difference b/w the means of two groups significant at 5% level of significance [$t_{0.05} = 2.086$ for 20 d.f].

→ Given: $n_1 = 12$ $\bar{x} = 74$ $\bar{x} = 74$ $s_1 = 8$

$n_2 = 10$ $\bar{y} = 70$ $\bar{y} = 70$ $s_2 = 10$

Hypothesis : There is no significant difference b/w population means of 2 groups.

WKT, $$t = \frac{\bar{x} - \bar{y}}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightarrow (*)$$

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s^2 = \frac{(11)(8)^2 + 9(10)^2}{12 + 10 - 2}$$

$$s^2 = 80.2$$

$$\boxed{s = 8.955445}$$

BOYS → 12 → n_1, 1 sample, $s_1 = 8$, $\bar{x} = 74$

GIRLS → 10 → n_2, 1 sample, $s_2 = 10$, $\bar{y} = 70$

(*) becomes,

$$t = \frac{\bar{x} - \bar{y}}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{74 - 70}{8.9554\sqrt{\frac{1}{12} + \frac{1}{10}}}$$

$$\boxed{t = 1.04316 < 2.086}$$

Our hypothesis is accepted

② Two types of batteries tested for their length of life & the following results are obtained. Compute student 't' & test whether there is a significant difference b/w two means. [$t_{0.05} = 2.101$ for 9 d.f]

→ Battery A : $n_1 = 10$ $\bar{x}_1 = 560$ hrs $\sigma_1^2 = 100 = s_1^2$

$n_2 = 10$ $\bar{x}_2 = 500$ hrs $\sigma_2^2 = 121 = s_2^2$

WKT, $$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightarrow (*)$$

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$= \frac{(10 - 1)(100) + (10 - 1)(121)}{10 + 10 - 2}$$

$$s^2 = 110.5$$

$$\boxed{s = 10.512}$$

(*) becomes, $$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{560 - 500}{10.512\sqrt{\frac{1}{10} + \frac{1}{10}}}$$

$$\boxed{t = 12.763 > 2.101}$$

Since our calculated t is greater than the tabulated t.

Our hypothesis is rejected

· TYPE 2 [S1 & S2 are not given] :-

(3) A Group of 10 Boys fed on Diet A & another group of 8 boys fed on different Diet B for period of 6 months recorded the following increase in weights :-

Diet A	5	6	8	1	12	4	3	9	6	10
Diet B	2	3	6	8	10	1	2	8		

Test whether diet's A & B differ significantly regarding their effect on increase in weights
[$t_{0.05} = 2.12$ for 16 d.o.f]

→ Let x denotes no. of boys who follows diet A,
y denotes no. of boys who follows diet B.

Hypothesis : There is no significant difference in diet A & diet B.

WKT, $t = \dfrac{\bar{x} - \bar{y}}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ → (*) where, $S^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$

$S^2 = \dfrac{\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2}{n_1+n_2-2}$

Consider, $s^2 = \dfrac{\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2}{n_1+n_2-2}$ → (*)

Now, $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{5+6+8+1+12+4+3+9+6+10}{10}$

$\boxed{\bar{x} = 6.4}$

$\Sigma(x-\bar{x})^2 = (5-6.4)^2 + (6-6.4)^2 + (8-6.4)^2 + (1-6.4)^2 + (12-6.4)^2 + (4-6.4)^2 + (3-6.4)^2 + (9-6.4)^2 + (6-6.4)^2 + (10-6.4)^2$

$\boxed{\Sigma(x-\bar{x})^2 = 102.4}$

Now, $\bar{y} = \dfrac{\Sigma y}{n} = \dfrac{2+3+6+8+10+1+2+8}{8}$

$\boxed{\bar{y} = 5}$

$\Sigma(y-\bar{y})^2 = (2-5)^2 + (3-5)^2 + (6-5)^2 + (8-5)^2 + (10-5)^2 + (1-5)^2 + (2-5)^2 + (8-5)^2$

$\boxed{\Sigma(y-\bar{y})^2 = 82}$

(1) becomes

$S^2 = \dfrac{(102.4) + 82}{10+8-2}$

$\boxed{S = 3.395}$

(*) becomes

$t = \dfrac{6.4-5}{3.395\sqrt{\frac{1}{10}+\frac{1}{8}}}$

$= 0.8693$

$t = 0.8963 < 2.12$

Hypothesis is accepted.

④ Two horses A & B were tested according to the time (in sec) to run a particular race with the following results.

Horse A : 28, 30, 32, 33, 33, 29, 34

Horse B : 29, 30, 30, 24, 27, 29

Test whether you can discriminate b/w two horses [$t_{0.05} = 2.2$ for 11 dof, $t_{0.02} = 2.72$ for 11 df].

→ Let x denotes Horse A

Let y denotes Horse B

Hypothesis :- There is no discrimination b/w 2 horses

$$\bar{x} = \frac{\Sigma x}{n} = \frac{28+30+32+33+33+29+34}{7} = 31.286$$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{29+30+30+24+27+29}{6} = 28.167$$

$$\Sigma(x-\bar{x})^2 = [(28-31.28)^2 + (30-31.28)^2 + (32-31.28)^2 + (33-31.28)^2 + (33-31.28)^2 + (29-31.28)^2 + (34-31.28)^2]$$

$$= \boxed{31.43}$$

$$\Sigma(y-\bar{y})^2 = [(29-28.167)^2 + 2(30-28.16)^2 + (24-28.16)^2 + (27-28.16)^2 + (29-28.16)^2]$$

$$= \boxed{26.83}$$

$$s^2 = \frac{1}{7+6-2}[31.43+26.83]$$

$$s^2 = 5.296$$

$$\boxed{s = 2.301}$$

$$t = \frac{\bar{x}-\bar{y}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{31.286-28.167}{2.301\sqrt{\frac{1}{7}+\frac{1}{6}}} = 2.43$$

Case 1 :- $t = 2.43 > t_{0.05} = 2.2$

Our hypothesis rejected at 5% level

Case 2 :- $t = 2.43 < t_{0.02} = 2.72$

Our hypothesis accepted at 2% level.

CHI-SQUARE TEST [χ^2 test]

WKT, $Z = \dfrac{\bar{x} - \mu}{\sigma}$

consider $Z^2 = \dfrac{(\bar{x} - \mu)^2}{\sigma^2} \rightarrow \chi^2$ for 1 df

$$Z^2 + Z^2 = \frac{(\bar{x}_1 - \mu_1)^2}{\sigma_1^2} + \frac{(\bar{x}_2 - \mu_2)^2}{\sigma_2^2} \rightarrow \chi^2 \text{ for 2 df}$$

$\vdots$

|||ly for n d.f $= \displaystyle\sum_{i=1}^{n} \frac{(x_i - \mu_i)^2}{\sigma_i^2}$

KARL has given for n-1 d.f

$$\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$$ where O_i - observed frequency, E_i - estimated frequency

① A die is thrown 264 times & the number appearing on the face, x follows the following frequency distribution.

x	1	2	4	5	6	3
f	40	30	56	52	60	26

Calculate the value of chi-square & test the hypothesis that the die is unbiased.

Given: $\left(\chi^{2(5)}_{0.03} = 11.07 \text{ \& } \chi^2_{0.01}(5) = 15.09\right)$

→ Expected no. of freq for getting $\{1,2,3,4,5,6\} = \dfrac{264}{6} = 44$

Hypothesis: Die is unbiased

x	1	2	4	5	6	3
observed frequency	40	30	56	52	60	26
estimated frequency	44	44	44	44	44	44

WKT, $\chi^2 = \Sigma \dfrac{(O_i - E_i)^2}{E_i}$

$$= \frac{(40-44)^2}{44} + \frac{(30-44)^2}{44} + \frac{(56-44)^2}{44} + \frac{(52-44)^2}{44} + \frac{(60-44)^2}{44} + \frac{(26-44)^2}{44}$$

$$\boxed{\chi^2 = 22.727}$$

case 1: $\chi^2 = 22.727 > \chi^2_{0.05}(5) = 11.07$

our hypothesis is rejected.

case 2: $\chi^2 = 22.727 > \chi^2_{0.01}(5) = 15.09$

our hypothesis is rejected.

② The following table gives the number of aircrafts accidents that occured during the various days of the week. Find whether the accidents are uniformly distributed over the week [Given $\chi^2_{0.05} = 12.59$ for 6 d.f]

Days	Sun	Mon	Tue	Wed	Thur	Fri	Sat
No. of accidents	14	16	8	12	11	9	14

→ observed frequency :- No. of accidents gives the observed frequency

estimated frequency :- $\frac{14+16+8+12+11+19+14}{7}$

$= 12$ accidents

Days	SUN	MON	TUE	WED	THUR	FRI	SAT
O_i	14	16	8	12	11	19	14
E_i	12	12	12	12	12	12	12

Hypothesis : accidents are uniformly distributed

WKT, $\chi^2 = \frac{\Sigma(O_i - E_i)^2}{E_i}$

$$= \frac{(14-12)^2}{12} + \frac{(16-12)^2}{12} + \frac{(8-12)^2}{12} + \frac{(12-12)^2}{12} + \frac{(11-12)^2}{12} + \frac{(19-12)^2}{12} + \frac{(14-12)^2}{12}$$

$$\boxed{\chi^2 = 7.5}$$

$\chi^2 = 7.5 < \chi^2_{0.05} = 12.59$

∴ our hypothesis is accepted

③ Sample analysis of examination results of 500 students was made. It was found that 220 students had failed, 170 had secured third class, 90 had secured second class and 20 had secured first class. Do these figures support in general examination results which is in the ratio 4:3:2:1 for the respective categories.

[Given $\chi^2_{0.05} = 7.21$ for 3 df].

→ observed frequency :- Examination results gives the observed frequency

220, 170, 90, 20

estimated frequency :- Expected frequency are

4 : 3 : 2 : 1

40% 30% 20% 10%

$\frac{40}{100}$ $\frac{30}{100}$ $\frac{20}{100}$ $\frac{10}{100}$

$\frac{4}{10}$ $\frac{3}{10}$ $\frac{2}{10}$ $\frac{1}{10}$

for 500 students,

$500\left(\frac{4}{10}\right)$ $500\left(\frac{3}{10}\right)$ $500\left(\frac{2}{10}\right)$ $500\left(\frac{1}{10}\right)$

200 150 100 50

O_i	220	170	90	20
E_i	200	150	100	50

Hypothesis : observed frequency will support the general frequency results ratio 4:3:2:1

WKT, $\chi^2 = \Sigma \frac{[O_i - E_i]^2}{E_i}$

$= \frac{(220-200)^2}{220} + \frac{(170-150)^2}{150} + \frac{(90-100)^2}{100} + \frac{(20-50)^2}{50}$

$\boxed{\chi^2 = 23.67}$

$\chi^2 = 23.67 > \chi^2_{0.05} = 7.21$ for 3 df.

TYPE 2 : [Best Fit for binomial distribution]

100%

① 4 coins are tossed 100 times & the following results are obtained. Fit a binomial distribution for the data & test the goodness of it.

No. of heads	0	1	2	3	4
Frequency	5	29	36	25	5

[$\chi^2_{0.05} = 9.44$ for 4 d.f]

→ Hypothesis : Binomial distribution is good fit to given data

observed frq : Given data gives observed frequency

Estimated frq : To calculate estimated frequency, we have to use,

$p(x) = {}^nC_x p^x q^{n-x}$, here $p, q = 1/2$, $n = 4$

$p(x) = {}^4C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{4-x}$

$p(x) = {}^4C_x \left(\frac{1}{2^x}\right)\left(\frac{1}{2^{4-x}}\right)$

$\boxed{p(x) = \frac{1}{16}\,{}^4C_x}$

$100p(x) = 100 * \frac{1}{16}\,{}^4C_x$ [∵ 100 trials]

$\boxed{f(x) = 6.25\,{}^4C_x}$

Let x denotes the no. of heads turning up

$f(0) = 6.25\,{}^4C_0 = 6.25*1 = 6.25 \approx 6$

$f(1) = 6.25\,{}^4C_1 = 6.25*4 = 25$

$f(2) = 6.25\,{}^4C_2 = 6.25*6 = 37.5 \approx 38$

$f(3) = 6.25\,{}^4C_3 = 6.25*4 = 25$

$f(4) = 6.25\,{}^4C_4 = 6.25*1 = 6.25$

O_i	5	29	36	25	5
E_i	6	25	38	25	6

$\Sigma O_i = \Sigma E_i$

WKT, $\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$

$= \frac{(5-6)^2}{6} + \frac{(29-25)^2}{25} + \frac{(36-38)^2}{38} + \frac{(25-25)^2}{25} + \frac{(5-6)^2}{6}$

$\boxed{\chi^2 = 1.078}$

$\chi^2 = 1.078 < \chi^2_{0.05} = 9.49$ [for 4 df]

Our hypothesis is accepted.

95% confidence that given data is good fit for binomial distribution

② Records taken, the no. of male & female birth in 800 families having 4 children as follows [$\chi^2_{0.05} = 9.49$ for 4 df]

No. of male births	0	1	2	3	4
No. of female births	4	3	2	1	0
No. of families	32	178	290	236	64

Test whether the data are consistent with the hypothesis that binomial law holds & chances of male birth is equal to female birth.

→ Hypothesis: Given Data is goodfit for Binomial distribution

Estimated frequency: $p(x) = {}^nC_x p^x q^{n-x}$ Here $p = q = 1/2$, $n = 4$

$p(x) = {}^4C_x (1/2)^x (1/2)^{4-x}$

$\boxed{p(x) = \frac{1}{16} {}^4C_x}$

$800 p(x) = 800 \frac{1}{16} {}^4C_x$

$\boxed{f(x) = 50 \, {}^4C_x}$

Let x denotes no. of male births,

$f(0) = 50 \, {}^4C_0 = 50$

$f(1) = 50 \, {}^4C_1 = 50 * 4 = 200$

$f(2) = 50 \, {}^4C_2 = 50 * 6 = 300$

$f(3) = 50 \, {}^4C_3 = 50 * 4 = 200$

$f(4) = 50 \, {}^4C_4 : 50 * 1 : 50$

Oi	32	178	290	236	64
Ei	50	200	300	200	50

$\Sigma O_i = \Sigma E_i$

WKT, $\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$

$= \frac{(32-50)^2}{50} + \frac{(178-200)^2}{200} + \frac{(290-300)^2}{300} + \frac{(236-200)^2}{200} + \frac{(64-50)^2}{50}$

$\boxed{\chi^2 = 19.633}$

$19.633 > 9.49$, ∴ our hypothesis is rejected.

③ A set of 5 similar coins is tossed 320 times & the result is

No. of heads	0	1	2	3	4	5
Frequency	6	27	72	112	71	32

Test the hypothesis that the data follow a binomial distribution.

→ (Similar to TYPE 2 1st problem)

100%.

④ Fit a poison distribution for the following data & check the goodness of it for 5% significance [$\chi_{0.05} = 9.49$ for 4 df]

x	0	1	2	3	4	5
f	110	170	130	60	23	7

→ Hypothesis: Poison fit is good fit for given data

observed freq: given data

estimated freq: estimated freq can be calculated by,

$p(x) = \frac{m^x e^{-m}}{x!}$

WKT, $\mu = \frac{\Sigma fx}{\Sigma f} = \cancel{110+170+130+60+23+7}$

$$= \frac{0+170+260+180+92+35}{500}$$

$$m = \boxed{\mu = 1.474}$$

$$p(x) = \frac{(1.474)^x e^{-1.474}}{x!}$$

$$500\, p(x) = 500\, \frac{(1.474)^x e^{-1.474}}{x!}$$

$$f(0) = \frac{500 \cdot e^{-1.474}}{0!} = 114.5 \approx 115$$

$$f(1) = \frac{500(1.474) \cdot e^{-1.474}}{1!} = 168.778 \approx 169$$

$$f(2) = \frac{500(1.474)^2 \cdot e^{-1.474}}{2!} = 125$$

$$f(3) = \frac{500(1.474)^3 \cdot e^{-1.474}}{3!} = 61$$

$$f(4) = \frac{500(1.474)^4 \cdot e^{-1.474}}{4!} = 23$$

$$f(5) = \frac{500(1.474)^5 \cdot e^{-1.474}}{5!} = 7$$

Oi	110	170	130	60	23	7
Ei	115	169	125	61	23	7

D.f = [No. of categories] - [no. of parameters estimated] - 1

= 6 - 1 - 1

= 4

WKT, $\chi^2 = \Sigma (O_i - E_i)^2 / E_i$

$$= \frac{(110-115)^2}{115} + \frac{(170-169)^2}{169} + \frac{(130-125)^2}{125} + \frac{(60-61)^2}{61} + \frac{(23-23)^2}{23} + \frac{(7-7)^2}{7}$$

$$\boxed{\chi^2 = 0.4337}$$

$\chi^2 = 0.4337 < \chi^2_{0.05} = 9.49$ (for 4df)

∴ Our hypothesis is accepted.

TYPE-3 :-

① Verify whether poison distribution can be assumed on the following data.

No. of defects	0	1	2	3	4	5
Frequencies	6	13	13	8	4	3

→ Hypothesis: Given data is good fit for poison distibution

observed frequency : 6 13 13 8 4 3

estimated frequency : $p(x) = \frac{m^x e^{-m}}{x!}$

$$\mu = \frac{\Sigma fx}{\Sigma f} = \frac{0+13+26+24+16+15}{47} = \frac{94}{47} = 2$$

$$\boxed{m = \mu = 2}$$

$$p(x) = \frac{2^x e^{-2}}{x!}$$

$$47\, p(x) = 47\, \frac{2^x \cdot e^{-2}}{x!}$$

$$\boxed{f(x) = 47\, \frac{2^x \cdot e^{-2}}{x!}}$$

Let x denotes no. of defective items,

$f(0) = 47 \frac{2^0 e^{-2}}{0!} = 6.36 \approx 6$

$f(1) = 47 \frac{2^1 e^{-2}}{1!} : 12.72 \approx 13$

$f(2) = 47 \frac{2^2 e^{-2}}{2!} : 12.72 \approx 13$

$f(3) = 47 \frac{2^3 e^{-2}}{3!} = 8.48 \approx 9$

$f(4) = 47 \frac{2^4 e^{-2}}{4!} : 4.24 \approx 4$

$f(5) = 47 \frac{2^5 e^{-2}}{5!} : 1.69 \approx 2$

O_i	6	13	13	8	4	3
E_i	6	13	13	9	4	2

$\Sigma O_i = \Sigma E_i$ & $E_i \geq 5$

$\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$

O_i	6	13	13	8	7
E_i	6	13	13	9	6

$= \frac{(6-6)^2}{6} + \frac{(13-13)^2}{13} +$

$\frac{(13-13)^2}{13} + \frac{(8-9)^2}{9} + \frac{(7-6)^2}{6}$

$\boxed{\chi^2 = 0.277}$

$(\chi^2_{0.05})$

$0.277 < 7.815$ for 3 d.f

$\therefore$ Our hypothesis is accepted

100% 8M

② 5 dice were thrown 96 times & the numbers 1,2,3 appearing on the face of dice are as follows:-

Test the hypothesis if it follows Binomial distribution.

No. of dice appearing 1,2,3	5	4	3	2	1	0
Frequency	7	19	35	24	8	3

→ observed frequency : 7 19 35 24 8 3

estimated frequency : $p(x) = {}^nC_x p^x q^{n-x}$ $\quad p = 1/2, q = 1/2$, $n = 5$

$= {}^5C_x (1/2)^x (1/2)^{5-x}$

$= {}^5C_x \frac{1}{2^x} \frac{1}{2^{5-x}}$

$= {}^5C_x \frac{1}{2^5}$

$\boxed{p(x) = \frac{1}{32} {}^5C_x}$

$96\, p(x) = 96 \frac{1}{32} {}^5C_x$

$\boxed{f(x) = 3\, {}^5C_x}$

Let x denotes no. of dice appearing 1,2 or 3

$f(5) = 3\, {}^5C_5 = 3 \times 1 = 3$

$f(4) = 3\, {}^5C_4 = 3 \times 5 = 15$

$f(3) = 3 \cdot {}^5C_3 = 3 \times 10 = 30$

$f(2) = 3 \cdot {}^5C_2 = 3 \times 10 = 30$

$f(1) = 3 \cdot {}^5C_1 = 3 \times 5 = 15$

$f(0) = 3 \cdot {}^5C_0 = 3 \times 1 = 3$

Oi	7	19	35	24	8	3
Ei	3	15	30	30	15	3

$\Sigma Oi = \Sigma Ei$

$Ei \geq 5$

we hv used 2 conditions

6-1-2 = 3 d.f

Oi	26	35	24	11
Ei	18	30	30	18

WKT, $\chi^2 = \Sigma \frac{(Oi-Ei)^2}{Ei}$

$= \frac{(26-18)^2}{18} + \frac{(35-30)^2}{30} + \frac{(24-30)^2}{30} + \frac{(11-18)^2}{18}$

$\boxed{\chi^2 = 8.311}$

$\chi^2 = 8.311 > \chi^2 = 7.815$ (for 3 d f)

Our hypothesis is rejected

*F-Test :- It is used to find equality of 2 population variance. we need to do f-test for testing whether 2 samples are came from the population which has equal variance.

$F = \frac{S_1^2}{S_2^2}$ where $S_1^2 > S_2^2$ $S_1^2 = \frac{\Sigma(x-\bar{x})^2}{n_1 - 1}$ $\bar{x} = \frac{1}{N}\Sigma x$

$F : \frac{S_2^2}{S_1^2}$ where $S_2^2 > S_1^2$ $S_2^2 = \frac{\Sigma(\cancel{y}-\bar{y})^2}{n_2 - 1}$ $\bar{y} = \frac{1}{N}\Sigma y$

TYPE-1

① The time taken by workers in performing a job by method I & method II is given below.

Method I	20	16	26	27	23	22	
Method II	27	33	42	35	32	34	38

Does the data show that the variance of time distribution from population from which these samples are drawn do not differ significantly F(6,5) at 0.05 = 4.96.

→ There is no significant difference between population means → Hypothesis

Method 1: 6 → n_1, 1 sample

Method 2: 7 → n_2, 1 sample

x	y	$x-\bar{x}$	$(x-\bar{x})^2$	$y-\bar{y}$	$(y-\bar{y})^2$
20	27	-2.3	5.29	-7.4	54.76
16	33	-6.3	39.69	-1.4	1.96
26	42	3.7	13.69	7.6	57.76
27	35	4.7	22.09	0.6	0.36
23	32	0.7	0.49	-2.4	5.76
22	34	-0.3	0.09	-0.4	0.16
	38			3.6	12.96

$\Sigma x = 134$ $\Sigma y = 241$ $\Sigma(x-\bar{x})^2 = 81.34$ $\Sigma(y-\bar{y})^2 = 133.72$

$\bar{x} = \frac{1}{n}\Sigma x = \frac{1}{6}(134) = 22.33$

$\bar{y} = \frac{1}{n}\Sigma y = \frac{1}{7}(241) = 34.43$

Consider $S_1^2 = \Sigma(x-\bar{x})^2/n_1 - 1$

$= 81.34/6-1$

$\boxed{S_1^2 = 16.26}$

Consider $S_2^2 = \Sigma(y-\bar{y})^2/n_2 - 1$

$= 133.72/7-1$

$\boxed{S_2^2 = 22.29}$

Here $S_2^2 > S_1^2$, $F = S_2^2/S_1^2 = 22.29/16.26 = \boxed{1.37}$

$\therefore F = 1.37 < F_{(6,5)}$ at 0.05 = 4.96

Our hypothesis is accepted

Sl. No.	Date	Title	Page No.	Teacher Sign/ Remarks

② A plant has installed two machines producing polythene bags. During the installation, the manufacturer of the machine has started that the capacity of the machine is to produce 20 bags in a day.

During to various factors such as different operators working on these machines, raw material etc, there is a variation in the number of bags produced at the end of the day.

The company researcher has taken a random sample of bags produced in 10 days for machine I and 13 days for machine II respectively.

The following data gives the no. of units of an item produced on a sampled day by the two machines.

Method I	20	16	26	27	23	22	18	24	25	19		
Machine II	27	33	42	35	32	34	38	28	41	43	30	37

Is there any significant difference b/w two variance

[$F(11,9) = 3.16$]

→ $\boxed{\text{Machine I}}$ — $10 \rightarrow n_1$ sample

$\boxed{\text{Machine II}}$ — $12 \rightarrow n_2$ sample

x	y	$x-\bar{x}$	$(x-\bar{x})^2$	$y-\bar{y}$	$(y-\bar{y})^2$
20	27	-2	4	-8	64
16	33	-6	36	-2	4
26	42	4	16	7	21
27	35	5	25	0	0
23	32	1	1	-3	9
22	34	0	0	-1	1
18	38	-4	16	3	9
24	28	2	4	-7	49
25	41	3	9	6	36
19	43	-3	9	8	64
	30			-5	25
	37			2	4

$\Sigma x = 220$ $\Sigma y = 420$ $\Sigma(x-\bar{x})^2 = 120$ $\Sigma(y-\bar{y})^2 = 314$

Consider $S_1^2 = \frac{\Sigma(x-\bar{x})^2}{n_1 - 1}$

$= \frac{120}{10-1}$

$\boxed{S_1^2 = 13.33}$

Consider $S_2^2 = \frac{\Sigma(y-\bar{y})^2}{n_2 - 1}$

$= \frac{314}{12-1}$

$\boxed{S_2^2 = 28.55}$

Here $S_2^2 > S_1^2$

$F = \frac{S_2^2}{S_1^2}$

$= \frac{28.55}{13.33} = 2.142$

$\boxed{F = 2.142}$

$F = 2.142 < F = 3.16$

Our hypothesis is accepted

③ In a sample of 8 observation, the sum of squares of deviation from the mean is 94.5.
In other sample of 10 observations the sum of squares of deviation from the mean is 101.7.
Test whether is there any significant differeence of variance $[F_{0.05}(7,9) = 3.29]$

→ $n_1 = 8 \quad \Sigma(x-\bar{x})^2 = 94.5$

$n_2 = 10 \quad \Sigma(Y-\bar{Y})^2 = 101.7$

H: there is no significant difference b/w the variance

WKT $F = \frac{S_1^2}{S_2^2}$ where $S_1^2 > S_2^2$

$F = \frac{S_2^2}{S_1^2}$ where $S_2^2 > S_1^2$

→ Consider

$S_1^2 = \frac{1}{n_1 - 1} \Sigma(x-\bar{x})^2$

$= \frac{1}{7}(94.5)$

$\boxed{S_1^2 = 13.5}$

→ Consider

$S_2^2 = \frac{1}{n_2 - 1} \Sigma(Y-\bar{Y})^2$

$= \frac{1}{9}(101.7)$

$\boxed{S_2^2 = 11.3}$

Here $S_1^2 > S_2^2$

$F = \frac{S_1^2}{S_2^2} = \frac{13.5}{11.3} = 1.195$

$F = 1.195 < F_{0.05}(7,9) = 3.29$

∴ our hypothesis is accepted.

TYPE - 2

***

① A normal population has two parameters mean (μ) & variance (σ^2). To test if two independent samples have been drawn from the same normal population

[$F_{0.05}(9,11) = 2.90$, $t_{0.05}$ for 20 df $= 2086$]

Sample	Size	Sample Mean	Sum of square of deviation from Mean
1	10	15	90
2	12	14	108

Test whether the samples come from the same normal population.

→ $n_1 = 10$ $\quad \bar{x} = 15$ $\quad \Sigma(x-\bar{x})^2 = 90$

$n_2 = 12$ $\quad \bar{y} = 14$ $\quad \Sigma(y-\bar{y})^2 = 108$

Hypothesis: There is no significant difference b/w variance. $\sigma_x^2 = \sigma_y^2$

Case 1: F Test

Consider, $s_1^2 = \frac{\Sigma(x-\bar{x})^2}{n_1-1}$

$= \frac{90}{10-1}$

$\boxed{s_1^2 = 10}$

Consider, $s_2^2 = \frac{\Sigma(y-\bar{y})^2}{n_2-1}$

$= \frac{108}{12-1}$

$\boxed{s_2^2 = 9.82}$

Here $s_1^2 > s_2^2$

WKT, $F = \frac{s_1^2}{s_2^2} = \frac{10}{9.82} = \boxed{1.02}$

∴ $F = 1.02 < 2.90$

our hypothesis is accepted

Case 2: t test

Hypothesis: There is no significant diff b/w population means.

i.e., $\mu_x = \mu_y$

WKT, $t = \frac{\bar{x}-\bar{y}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ → (*) where $s^2 = \frac{\Sigma(x-\bar{x})^2 + (y-\bar{y})^2}{n_1+n_2-2}$

$= \frac{90+108}{10+12-2}$

$\boxed{s = 3.14}$

$= \frac{15-14}{3.14\sqrt{\frac{1}{10}+\frac{1}{12}}}$

$\boxed{t = 0.742}$

∴ $t = 0.742 < t_{0.05} = 2.086$

∴ our hypothesis is accepted

Here we observed that,

case 1: $\sigma_x^2 = \sigma_y^2$

case 2: $\mu_x = \mu_y$

∴ Both the population are same

Hence given sample comes from the sample population.

JOINT PROBABILITY

→ If x & y are two discrete random variables, we define joint probability function x & y by

$$P(x=x, y=y) = F(x,y)$$

→ A function $F(x,y)$ is said to be probability mass function if it satisfies the following conditions :-

- $F(x,y) > 0$
- $\sum_x \sum_y F(x,y) = 1$

JOINT PROBABILITY TABLE

Let $X = \{x_1, x_2, \ldots\ldots x_n\}$

$Y = \{y_1, y_2 \ldots\ldots y_n\}$

$P(x_i\ y_i) = J_{ij}$

x \ y	y_1	y_2	....	y_n	sum
x_1	J_{11}	J_{12}	....	J_{mn}	$f(x_1)$
x_2	J_{21}	J_{22}	....	J_{2n}	$f(x_2)$
⋮				⋮	⋮
x_m	J_{m1}	J_{m2}	....	J_{mn}	$f x m$
	$g(y_1)$	$g(y_2)$	....	$g(n)$	

$\{f(x_1)\ f(x_2)\ f(x_3) \ldots\ldots f(x_m)\}$

$\{g(y_1)\ \ g(y_2)\ \ g(y_3) \ldots\ldots g(y_n)\}$

Expectation :-

$\mu_x = E(x) = \Sigma x \cdot f(x)$

$\mu_y = E(y) = \Sigma y \cdot g(y)$

$E(xy) = \Sigma x_i y_j J_{ij}$

Variance :-

$\sigma_x^2 = E(x^2) - \mu_x^2$, where $E(x^2) = \Sigma x^2 \cdot f(x)$

$\sigma_y^2 = E(y^2) - \mu_y^2$, where $E(y^2) = \Sigma y^2 \cdot g(y)$

Co-variance :-

$Cov(x,y) = E(xy) - E(x)E(y)$

Correction :-

$$\rho(xy) = \frac{Cov(xy)}{\sigma_x \sigma_y}$$

NUMERICALS :-

10M ***

① The joint distribution of two random variables

x \ y	-2	-1	4	5
1	0.1	0.2	0	0.3
2	0.2	0.1	0.1	0

Determine the marginal probability,

(a) $E(x)$, $E(y)$ & $E(x,y)$

(b) Standard deviation of x, y

(c) Covariance of x & y

(d) Corelation of x & y

Further verify that x & y are dependent random variable.

→ Marginal distribution of x

x	1	2
$f(x)$	0.6	0.4

Marginal distribution of y

y	-2	-1	4	5
$g(y)$	0.3	0.3	0.1	0.3

- $E(x) = \Sigma(x \cdot f(x))$

$= 1(0.6) + 2(0.4)$

$\boxed{E(x) = 1.4}$

- $E(y) = \Sigma(y \cdot g(y))$

$= -2(0.3) - 1(0.3) + 4(0.1) + 5(0.3)$

$\boxed{E(y) = 1}$

$E(xy) = \Sigma(xy J_{ij})$

$= (1)(-2)(0.1) + (1)(-1)(0.2) + (1)(4)(0) + (1)(5)(0.3)$
$+ (2)(-2)(0.2) + 2(-1)(0.1) + (2)(4)(0.1) + 2(5)(0)$

$= \boxed{0.9}$

• Variance :-

$\sigma_x^2 = E(x^2) - \mu_x^2 \rightarrow ①$

Consider, $E(x^2) = \Sigma(x^2 . f(x))$

$= 1^2(0.6) + 2^2(0.4)$

$= 0.6 + 1.6$

$\boxed{E(x^2) = 2.2}$

① becomes,

$\sigma_x^2 = 2.2 - (1.4)^2$

$= 2.2 - 1.96$

$= 0.24$

$\boxed{\sigma_x = \sqrt{0.24}}$

$\sigma_y^2 = E(y^2) - \mu_y^2 \rightarrow ②$

Consider, $E(y^2) = \Sigma(y^2 . g(y))$

$= (-2)^2(0.3) + (-1)^2(0.3)$

$+ 4^2(0.1) + 5^2(0.3)$

$\boxed{E(y^2) = 10.6}$

② becomes,

$\sigma_y^2 = 10.6 - (1)^2$

$\sigma_y^2 = 9.6$

$\boxed{\sigma_y = \sqrt{9.6}}$

• $Cov(xy) = E(xy) - E(x).E(y)$

$= 0.9 - (1.4)(1)$

$= 0.9 - 1.4$

$\boxed{Cov(xy) = -0.5}$

Correlation, (x,y) $\rho(x,y) = \dfrac{Cov(xy)}{\sigma_x . \sigma_y}$

$= \dfrac{-0.5}{(0.49)(3.1)}$

$\boxed{\rho(x,y) = -0.33}$

② NOTE :- How to verify that x & y are dependable random variables.

$\boxed{f(x_1) = g(y_2) = J_{ij}}$

② The Joint Probability Distribution of 2 random variables x & y is as follows:-

x \ y	-4	2	7
1	1/8	1/4	1/8
5	1/4	1/8	1/8

(a) Determine marginal distribution of x & y.
(b) E(x) & E(y)
(c) E(xy)
(d) σ_x & σ_y
(e) Cov(xy)
(f) $\rho(xy)$
(g) verify that x & y are dependent.

→

x \ y	-4	2	7	f(x)
1	1/8	1/4	1/8	4/8
5	1/4	1/8	1/8	4/8
g(y)	3/8	3/8	2/8	1

Marginal Distribution of x

x	1	5
f(x)	4/8	4/8

Marginal Distribution of y

y	-4	2	7
g(y)	3/8	3/8	2/8

$E(X) = \Sigma(x \cdot f(x))$

$= 1\left(\frac{4}{8}\right) + 5\left(\frac{4}{8}\right)$

$= 4/8 + 20/8$

$= \boxed{3}$

$E(Y) = \Sigma(Y \cdot g(Y))$

$= -4(3/8) + 2(3/8) + 7(2/8)$

$= -\frac{12}{8} + \frac{6}{8} + \frac{14}{8}$

$= \boxed{1}$

$E(XY) = \Sigma(x_i y_i J_{ij})$

$= (1)(-4)(1/8) + (1)(2)(1/4) + (1)(7)(1/8)$

$+ (5)(-4)(1/4) + 5(2)(1/8) + 5(7)(1/8)$

$= -\cancel{1/2} + 1/2 + 7/8 - 5 + 5/4 + 35/8$

$= 12/8$

$= \boxed{3/2}$

• $\sigma_x^2 = E(X^2) - \mu_x^2 \rightarrow (1)$

Consider $E(X^2) = \Sigma(x^2 f(x))$

$= 1^2(4/8) + 5^2(4/8)$

$= 4/8 + 100/8$

$= \boxed{13}$

$\sigma_Y^2 = E(Y^2) - \mu_Y^2 \rightarrow (2)$

Consider $E(Y^2) \cdot g(Y)$

$= (-4)^2(3/8) + (2)^2(3/8) + 7^2(2/8)$

$= 48/8 + 12/8 + 98/8$

$= \boxed{79/4}$

(1) becomes,

$\sigma_x^2 = 13 - (3)^2 = \boxed{4}$

$\boxed{\sigma_x = 2}$

$\sigma_Y^2 = \frac{79}{4} - 1^2 = \frac{75}{4} \Rightarrow \boxed{\sigma_Y = 4.33}$

$Cov(XY) = E(XY) - E(X) \cdot E(Y)$

$= 3/2 - 3(1)$

$= \boxed{\frac{-3}{2}}$

• $\rho(XY) = \frac{Cov(XY)}{\sigma_x \cdot \sigma_Y}$

$= \frac{-3/2}{(2)(4.33)}$

$= \boxed{-0.1732}$

• WKT If X & Y are independent random variables then

$f(x_i) \cdot g(y_i) = J_{ij}$

$(4/8)(3/8) = 12/8 \neq J_{11}$

$[f(x_1)\ g(y_1) = J_{11}]$

$\therefore$ X & Y are dependent random variables

(3) X & Y are independent random variables, X takes the values 2, 5, 7 with probability 1/2, 1/4, 1/4. Y takes the values 3, 4, 5 with probability 1/3, 1/3, 1/3

(a) Find the joint probability that distribution of X & Y.

(b) Find covariance

(c) Find probability distribution of $Z = X + Y$.

→ Marginal distribution of X

x	2	5	7
$f(x)$	1/2	1/4	1/4

Marginal distribution of Y

Y	3	4	5
$g(Y)$	1/3	1/3	1/3

• Given X & Y are independent random variables, therefore we can construct the distribution table

X \ Y	3	4	5
2	1/6	1/6	1/6
5	1/12	1/12	1/12
7	1/12	1/12	1/12

$Cov(XY) = E(XY) - E(X)E(Y) \rightarrow ⊛$

$E(X) = \Sigma(x \cdot f(x))$

$= 2(1/2) + 5(1/4) + 7(1/4)$

$= \boxed{4}$

$E(Y) = \Sigma(y \cdot f(y))$

$= 3(1/3) + 4(1/3) + 5(1/3)$

$= \boxed{4}$

Consider, $E(XY) = \Sigma(x\,y\,J_{ij})$

$= (2)(3)(1/6) + 2(4)(1/6) + 2(5)(1/6)$
$+ 5(3)(1/12) + 5(4)(1/12) + 5(5)(1/12)$
$+ 7(3)(1/12) + 7(4)(1/12) + 7(5)(1/12)$

$= \boxed{16}$

⊛ becomes,

$Cov(XY) = E(XY) - E(X)E(Y)$

$= 16 - 4(4)$

$= \boxed{0}$

• we have to find the probability of $Z = X + Y$

$Z = 5, 6, 7, 8, 9, 10, 11, 12$

Z	5	6	7	8	9	10	11	12
f(z)	1/6	1/6	1/6	1/12	1/12	1/12	1/12	1/12

④ Suppose X & Y are independent random variable X takes the values 1, 2 with probability 0.7, 0.3 respectively. Y takes the values −2, 5, 8 with probability 0.3, 0.5, 0.2. Find

ⓐ Joint probability distribution of XY.

ⓑ Find cov(XY).

→ Given:-

x	1	2
f(x)	0.7	0.3

$f(x_1)$ $f(x_2)$

Marginal distribution of X

y	−2	5	8
g(y)	0.3	0.5	0.2

$g(y_1)$ $g(y_2)$ $g(y_3)$

Marginal distribution of Y

Given :- X & Y are independent random variable is $f(x_i) \cdot g(y_j) = J_{ij}$

X \ Y	−2	5	8
1	0.21	0.35	0.14
2	0.09	0.15	0.06

This is the required joint probability distribution

ⓑ $cov(XY) = E(XY) - E(X)E(Y) \rightarrow ⊛$

Consider, $E(X) = \Sigma x \cdot f(x)$

$= 1(0.7) + 2(0.3)$

$= \boxed{1.3}$

Consider $E(Y) = \Sigma Y f(Y)$

$= -2(0.3) + 5(0.5) + 8(0.2)$

$= \boxed{3.5}$

Consider $E(XY) = \Sigma X_i Y_j J_{ij}$

$= 1(-2)(0.21) + 1(5)(0.35) + 1(8)(0.14) + 2(-2)(0.09) + 2(5)(0.15) + 2(8)(0.06)$

$= \boxed{4.55}$

(*) becomes,

$cov(XY) = 4.55 - (1.3)(3.5)$

$= \boxed{0}$

TYPE-3 :-

① A fair coin tossed thrice & random variable X & Y are defined as follows :-

$X = 0$ or 1 accordingly as head or tails occurs on the first toss

Y = no. of heads

ⓐ determine distribution of X & Y.

ⓑ determine the joint probability distribution of X & Y

ⓒ Obtain the expectation of X, Y & XY. Also find standard deviations of X & Y.

ⓓ Compute co-variance & co-relation of X & Y

→ A fair coin tossed thrice :-

S	HHH	HHT	HTH	HTT	THH	THT	TTH	TTT
X	0	0	0	0	1	1	1	1
Y	3	2	2	1	2	1	1	0

(a) distribution of X & Y

$f(X=0) = \frac{4}{8} = \frac{1}{2}$

$f(X=1) = \frac{4}{8} = \frac{1}{2}$

$g(Y=0) = 1/8$

$g(Y=1) = 3/8$

$g(Y=2) = 3/8$

$g(Y=3) = 1/8$

x	0	1
$f(x)$	½	½

Y	0	1	2	3
g(y)	1/8	3/8	3/8	1/8

(b) Joint probability distribution :-

X \ Y	0	1	2	3
0	0	1/8	2/8	1/8
1	1/8	2/8	1/8	0

(c) $E(x) = \Sigma x f(x)$

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

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

$E(Y) = \Sigma Y g(Y)$

$= 0(\frac{1}{8}) + 1(\frac{3}{8}) + 2(\frac{3}{8}) + 3(\frac{1}{8})$

$= \boxed{3/2}$

$E(xy) = \Sigma x_i y_j J_{ij}$

$= 0(0)(0) + (0)(1)(\frac{1}{8}) + (0)(2)(\frac{2}{8}) + (0)(3)(\frac{1}{8}) + 1(0)(\frac{1}{8}) + (1)(1)(\frac{2}{8}) + (1)(2)(\frac{1}{8}) + 1(3)(0)$

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

$\sigma_x^2 = E(x^2) - \mu_x^2 \rightarrow$ ①

$\sigma_y^2 = E(Y)^2 - \mu_y^2 \rightarrow$ ②

* MORKOV CHAIN [Future Probability]

→ Probability Vector :-

A vector $V = V_1, V_2, V_3 \ldots V_n$ is said to be a prob vector.

i. Each one of its components are non - negative

ii. Sum of the components should be equal to 1

Ex:- i. $S = \{H, T\}$

$A = [1/2, 1/2]$

ii. $C = [0, 1]$

iii. $B = [1/4, 1/2, 1/4]$

→ STOCHASTIC MATRIX :-

It is a square matrix having every row in the form of Probability vector.

Ex:- $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{2\times2}$ $B = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}_{2\times2}$

→ REGULAR STOCHASTIC MATRIX :-

It is a regular stochastic matrix if all the entries of sum poweres of P^n are positive.

* NOTE :-

P has unique fixed point $V = V_1, V_2, V_3 \ldots V_n$ such that $VP = V$ where $V_1 + V_2 + \ldots V_n = 1$.

* TYPE 1 :-

① Find the unique fixed probability vector for regular stochastic matrix.

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1/6 & 1/2 & 1/3 \\ 0 & 2/3 & 1/3 \end{bmatrix}$$

→ We have to find $v = (x, y, z)$ where $x+y+z = 1$ such that $VA = v$.

$$[x \ y \ z] \begin{bmatrix} 0 & 1 & 0 \\ 1/6 & 1/2 & 1/3 \\ 0 & 2/3 & 1/3 \end{bmatrix} = [x \ y \ z]$$

$$\left[0 + y(1/6) + 0 \quad x + y/2 + \tfrac{2}{3}z \quad 0 + y/3 + z/3\right] = [x \ y \ z]$$

$$\left[y/6 \quad x + y/2 + \tfrac{2}{3}z \quad y/3 + z/3\right] = [x \ y \ z]$$

$\frac{y}{6} = x$

$y = 6x \rightarrow ①$

$x + y/2 + 2z/3 = y$

$6x - 3y + 4z = 0 \rightarrow ②$

$y/3 + z/3 = z$

$y - 2z = 0 \rightarrow ③$

we have,

$x + y + z = 1$

$x + 6x + z = 1$

$7x + z = 1$

$\boxed{z = 1 - 7x} \rightarrow ④$

$6x - 3(6x) + 4(1 - 7x) = 0$

$6x - 18x + 4 - 28x = 0$

$-40x + 4 = 0$

$x = 4/40$

$\boxed{x = 1/10}$

$y = 6 \times \frac{1}{10}$

$\boxed{y = 3/5}$

$y - 2z = 0$

$3/5 - 2z = 0$

$\boxed{z = \frac{3}{10}}$

$$[x \ y \ z] = [1/10 \ \ 3/5 \ \ 3/10]$$

---

② $$P = \begin{bmatrix} 0 & 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/4 & 1/4 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix}$$

→ we have to find $v = [a \ b \ c \ d]$ where $a+b+c+d = 1$ such that $vP = v$

$$[a \ b \ c \ d] \begin{bmatrix} 0 & 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/4 & 1/4 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix} = [a \ b \ c \ d]$$

$$\left[b/2 + c/2 + d/2, \ a/2 + c/2 + d/2, \ a/4 + b/4, \ a/4, b/4\right] = [a, b, c, d]$$

$b/2 + c/2 + d/2 = a$ | $a/2 + c/2 + d/2 = b$ | $a/4 + b/4 = c$ | $a/4 + b/4 = d$

we have,

$a + b + c + d = 1$

$a + 2a = 1$

$\boxed{a = 1/3}$

$a + b + c + d = 1$

$b + 2b = 1$

$\boxed{b = 1/3}$

$a + b = 4c$

$1/3 + 1/3 = 4c$

$\boxed{c = 1/6}$

$\therefore \boxed{d = 1/6}$

$$v = [a, b, c, d] = [1/3, 1/3, 1/6, 1/6]$$

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

→ we have to find $v = [a\ b]$ where $a+b=1$ such that $vA = v$

$$[a\ b] \begin{bmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{bmatrix} = [a\ b]$$

$$[3a/4 + b/2\ ,\ a/4 + b/2] = [a\ b]$$

$\frac{3a}{4} + b/2 = a$ | $a/4 + b/2 = b$

$3a + 2b = 4a \rightarrow ①$ | $a + 2b = 4b \rightarrow ②$

we have, $a+b=1$

$b = 1-a$

$3a + 2(1-a) = 4a$

$3a + 2 - 2a = 4a$

$a + 2 = 4a$

$\boxed{a = 3/2}$

$a + 2b = 4b$

$3/2 + 2b = 4b$

$3/2 = 2b$

$\boxed{b = 3/4}$

$v = [a, b] = [3/2, 3/4]$

---

* TYPE 2 :-

① S.T $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix}$ is regular stochastic matrix & also find associated unique fixed probability vector.

→ we have to show p is regular stochastic,

$$P^2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

$$P^2 = \begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$

$$P^3 = P^2 \cdot P = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$$

$$P^4 = P^3 \cdot P = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 1/4 & 1/2 \\ 1/4 & 1/2 & 1/4 \end{bmatrix}$$

$$P^5 = P^4 \cdot P = \begin{bmatrix} 1/4 & 1/4 & 1/2 \\ 1/4 & 1/2 & 1/4 \\ 1/8 & 3/8 & 1/2 \end{bmatrix}$$

we have to find unique fixed probability vector $v(x, y, z)$ where $x+y+z=1$

$$[x\ y\ z] \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix} = [x\ y\ z]$$

$$[z/2\ ,\ x + z/2\ ,\ y] = [x\ y\ z]$$

$z/2 = x$, $x + z/2 = y$, $y = z$

$x - z/2 = 0$

$2x - z = 0$

$x - y + z/2 = 0$

$2x - 2y + z = 0$

$y - z = 0$

we have, $x + y + z = 1$

$y = 1 - x - z$

$1 - x - z - z = 0$

$1 - x - 2z = 0$

$x + 2z = 1$

$x + 4x = 1$

$\boxed{x = 1/5}$

$2(1/5) - z = 0$

$\boxed{z = \frac{2}{5}}$

$[x \ y \ z] = [1/5 \ \ 2/5 \ \ 2/5]$

---

* TYPE 3 :-

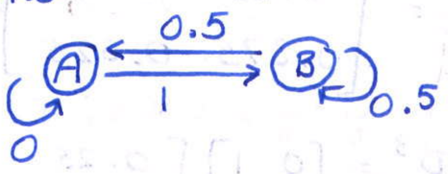
① A habitual gambler is a member of two clubs A & B. He visits either of the clubs everyday for playing cards. He never visits club A on two consecutive days. But, if he visits club on a partiadar day, then the next day he is as likely to visit club B, or club A. Find the transition matrix of this markov chain. Also

(a) Show that matrix is a regular stochastic matrix & find unique fixed probability vector.

(b) If the person has visited club B on monday, find the probability that he visits club A on thursday.

→ Transition probability matrix

$$P = \begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{bmatrix} 0 & 1 \\ 0.5 & 0.5 \end{bmatrix} \end{array}$$



(a) Consider : $P^2 = P.P = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}$

we observe that in P^2, all entries are positive

$\therefore$ P is a regular stochastic matrix

we have to find unique fixed probability vector $v(x, y)$ where $x + y = 1$ such that $VP = V$

$$[x \ y] \begin{bmatrix} 0 & 1 \\ 0.5 & 0.5 \end{bmatrix} = [x \ y]$$

$$[0.5y \ , \ x + 0.5y] = [x \ y]$$

$y(0.5) = x$

$x = 0.5y$

$x + 0.5y = y$

$x = 0.5y$

$x + y = 1$

$0.5y + y = 1$

$1/2 y + y = 1$

$\boxed{y = 2/3}$

$x = 1/2 y$

$\boxed{x = 1/3}$

$v(x,y) = [1/3 \ \ 2/3]$

(b) Let us suppose Monday as day 1 then thursday will be 3 days after monday

we have to find $[0 \ \ 1] P^3$

$$P^2 = P.P = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}$$

$$P^3 = \begin{bmatrix} 0.25 & 0.75 \\ 0.375 & 0.625 \end{bmatrix}$$

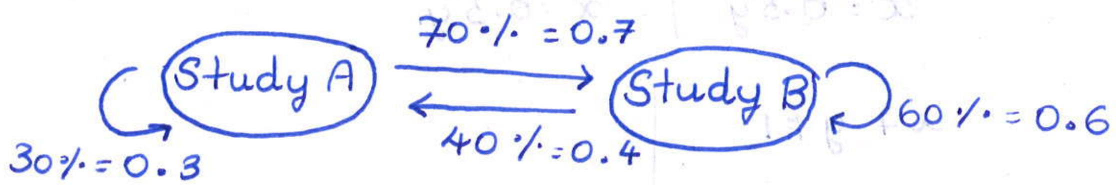
$$VP^3 = [0 \quad 1] \begin{bmatrix} 0.25 & 0.75 \\ 0.375 & 0.625 \end{bmatrix}$$

$$VP^3 = [0.375 \quad 0.625]$$

probability that he visits club A on thursday is 0.375.

② A student's study habits are as follows. If he studies one night, he is 70% sure not to study the next night. On the other hand if he does not study one night, he is 60% sure not to study the next night. In the long run how often does he study.

→



Transition probability matrix (t.p.m)

$$P = \begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{bmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{bmatrix} \end{array}$$

In order to find in the long ron, we have to find unique fixed probability vector $V(x,y)$ where $x+y=1$ such that $VP = V$

$$[x \quad y] \begin{bmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{bmatrix} = [x \quad y]$$

$$[0.3x + 0.4y \quad , \quad 0.7x + 0.6y] = [x \quad y]$$

$0.3x + 0.4y = x$

$0.3x - x + 0.4y = 0$

$-0.7x + 0.4y = 0$

$0.7x + 0.6y = y$

$0.7x - 0.4y = 0$

we have,

$x + y = 1$

$y = 1 - x$

$0.7x - 0.4(1-x) = 0$

$0.7x - 0.4 + 0.4x = 0$

$\boxed{x = 4/11}$

$0.7 \times \frac{4}{11} = 0.4y$

$\boxed{y = 7/11}$

$V(x,y) = (4/11, 7/11)$

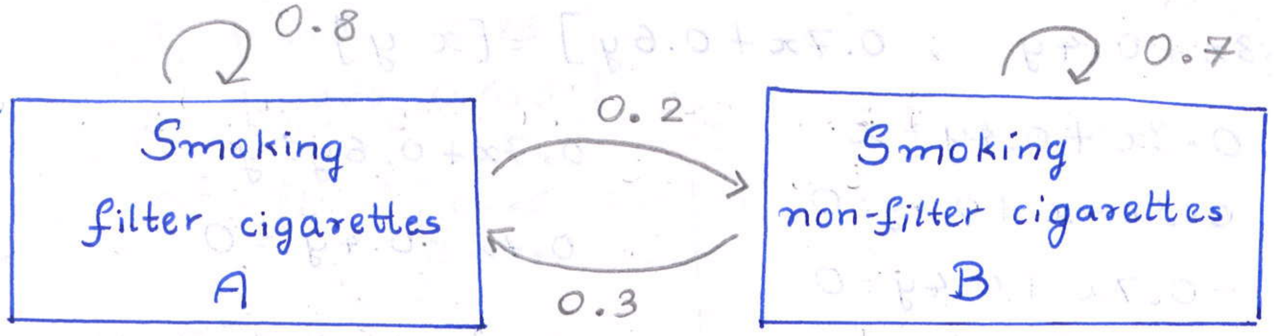
study doesn't study

$x = \frac{4}{11} \times 100 = 36.36\%$

$y = \frac{7}{11} \times 100 = 63.63\%$

---

③ A man's smoking habits are as follows. If he smokes filter cigarattes one week, he switches to non-filter cigarattes the next week with probability 0.2. On the other hand, if he smoke non-filter cigaratte one week, there is a probability of 0.7 that he will smoke non-filter cigarattes the next week as well. In the long run how often does he smoke filter cigarattes.



Transition probability matrix

$$P = \begin{array}{c} A \\ B \end{array} \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}$$

To find long run probability we should find unique fixed probability vector.

Now unique fixed probability vector $v(x,y)$ where $x+y=1$ such that

$$vP = v$$

$$[x \quad y]\begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} = [x \quad y]$$

$$[0.8x + 0.3y \quad , \quad 0.2x + 0.7y] = [x \quad y]$$

$0.8x + 0.3y = x$

$-0.2x + 0.3y = 0$

$0.2x + 0.7y = y$

$0.2x - 0.3y = 0$

we have,

$x + y = 1$

$x = 1 - y$

$0.2(1-y) - 0.3y = 0$

$0.2 - 0.2y - 0.3y = 0$

$0.2 - 0.5y = 0$

$\boxed{y = 2/5}$ $\boxed{x = 3/5}$ $V(x,y) = [3/5 \quad 2/5]$

filter cigarettes — non filter cigarettes

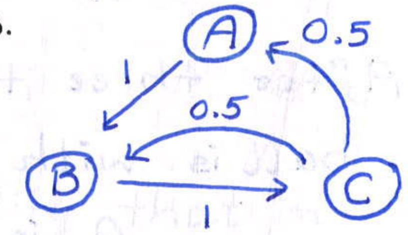
3X3 MATRIX

④ Three Boys A, B, C are throwing ball to each other, A always throws the ball to B and B always throws the ball to C. C is just as likely to throw the ball to B as to A. If C was the first person to throw the ball, find the probability that after 3 throws.

(i) A has the ball

(ii) B has the ball

(iii) C has the ball



→ Transition probability matrix,

$$P = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.5 & 0.5 & 0 \end{bmatrix} \end{array}$$

Given, C was the first person to throw the ball i.e., $[0 \quad 0 \quad 1]$

After 3 throws we have to find $[0 \quad 0 \quad 1]P^3$

Consider $P^2 = P \times P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.5 & 0.5 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.5 & 0.5 & 0 \end{bmatrix}$

$$P^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \end{bmatrix}$$

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

Consider (*) $[0 \quad 0 \quad 1]P^3$

$$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

$$= \begin{bmatrix} \underset{A}{0.25} & \underset{B}{0.25} & \underset{C}{0.5} \end{bmatrix}$$

After three throws the probability that the ball is with

A is 0.25

B is 0.25

C is 0.5

---

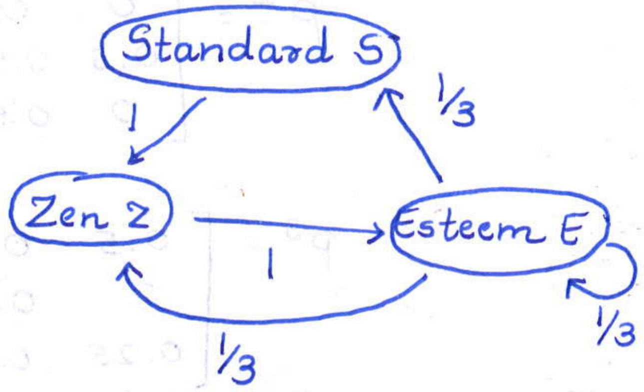
(5) Each year a man trades his car for a new car in 3 brands of the popular company Maruti Udyog limited. If he has a 'standard' he trades it for 'zen'. If he has a 'zen' he trades it for a 'esteem'. If he has a 'esteem' he is just likely to trade it for a new 'esteem' or for a 'zen' or a standard one. In 1996 he bought his first car which was esteem. Find the probability that he has

(i) (a) 1998 esteem

(b) 1998 standard

(c) 1999 zen

(d) 1999 esteem



(ii) In the long run, how often will he have a esteem?

→ Transition probability matrix

$$P = \begin{array}{c} \\ S \\ Z \\ E \end{array} \begin{array}{c} \begin{array}{ccc} S & Z & E \end{array} \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} \end{array}$$

Given: In 1996 he bought his first car which was esteem i.e., $[0 \ 0 \ 1]$

(i) (a) 1998 esteem

we have to find the probability that in 1998 he takes the esteem and standard

1996, after 2 years = 1998

we have to find: $[0 \ 0 \ 1] P^2$

$$P^2 = P.P = \begin{bmatrix} 0 & 0 & 1 \\ 0.333 & 0.333 & 0.333 \\ 0.111 & 0.444 & 0.444 \end{bmatrix}$$

$$[0 \ 0 \ 1] \begin{bmatrix} 0 & 0 & 1 \\ 0.333 & 0.333 & 0.333 \\ 0.111 & 0.444 & 0.444 \end{bmatrix}$$

$$[0.111 \quad 0.444 \quad 0.444]$$

$$\begin{bmatrix} \underset{\text{standard}}{1/9} & \underset{\text{zen}}{4/9} & \underset{\text{esteem}}{4/9} \end{bmatrix}$$

(b) 0.111 or 1/9 probability that he'll have standard in 1998

(c) (d) we have to find the probability that he has zen / esteem in 1999 :-

1996, after 3 years

$= [0 \quad 0 \quad 1] P^3$

$$P^3 = P^2 \cdot P = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/9 & 4/9 & 4/9 \\ 4/27 & 7/27 & 16/27 \end{bmatrix}$$

$[0 \quad 0 \quad 1] * P^3$

$$= \begin{bmatrix} \underset{\text{standard}}{\underset{\downarrow}{4/27}} & \underset{\text{zen}}{\underset{\downarrow}{7/27}} & \underset{\text{esteem}}{\underset{\downarrow}{16/27}} \end{bmatrix}$$

(ii) To find long run probability of esteem, we have to find unique fixed probability vector $V = (x, y, z)$ when $x+y+z=1$ such that $vP = v$

$$[x \ y \ z] \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} = [x \ y \ z]$$

$$\left[z/3 \quad x + z/3 \quad y + z/3 \right] = [x \ y \ z]$$

$z/3 = x$	$x + z/3 = y$	$y + z/3 = z$
$3x - z = 0$	$3x - 3y + z = 0$	$3y - 2z = 0$

we have,

$x+y+z=1$

$x = 1-y-z$

$3x - z = 0$

$3(1-y-z) - z = 0$

$3 - 3y - 3z - z = 0$

$-3y - 4z = -3 \rightarrow (1)$

$3x - 3y + z = 0 \rightarrow (2)$

$3y - 2z = 0 \rightarrow (3)$

Upon solving,

$\boxed{x = 1/6}$ $\boxed{y = 1/3}$ $\boxed{z = 1/2}$

$$v(x, y, z) = \begin{bmatrix} \underset{\text{standard}}{1/6} & \underset{\text{zen}}{1/3} & \underset{\text{esteem}}{1/2} \end{bmatrix}$$

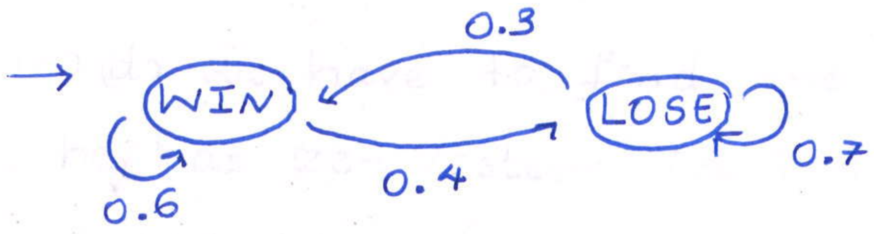
---

(5) A gambler's luck follows a pattern. If he wins a game the probability of winning the next game is 0.6. However if he loses a game, the probability of losing the next game is 0.7. There is an even chance of gambler winning the game. If so,

(a) What is the probability of he winning the 2nd game?

(b) What is the probability of he winning the 3rd game?

(c) In the long run how often he will win?



$$P = \begin{array}{c|cc} & W & L \\ W & 0.6 & 0.4 \\ L & 0.3 & 0.7 \end{array}$$

(a) probability of winning the first game is $1/2$

initial probability vector $p = [0.5 \quad 0.5]$

Consider $p^{(0)} \cdot P$

$$= [0.5 \quad 0.5]\begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix}$$

$$= [0.45 \quad 0.55]$$

or

$$\left[\underset{\text{win}}{9/20} \quad \underset{\text{lose}}{11/20}\right]$$

probability that he'll win the 2^{nd} game is $9/20$

(b) Consider, $p^{(0)} p^{(2)}$

$$P^2 = \begin{bmatrix} 0.48 & 0.52 \\ 0.39 & 0.61 \end{bmatrix}$$

$$[0.5 \quad 0.5]\begin{bmatrix} 0.48 & 0.52 \\ 0.39 & 0.61 \end{bmatrix}$$

$$= [0.435 \quad 0.565]$$

$$= \left[\underset{W}{87/200} \quad \underset{L}{113/200}\right]$$

probability that he'll win 3^{rd} game is $87/200$

(c) we shall find the fixed probability vector $v(x,y)$ where $x+y=1$, such that $vp = v$

$$[x \quad y]\begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix} = [x \quad y]$$

$0.6x + 0.3y = x$	$0.4x + 0.7y = y$
$-0.4x + 0.3y = 0$	$0.4x - 0.3y = 0$

we have, $x + y = 1$

$x = 1 - y$

$-0.4(1-y) + 0.3y = 0$

$-0.4 + 0.4y + 0.3y = 0$

$-0.4 + 0.7y = 0$

$y = \frac{0.4}{0.7}$ $\qquad$ $x = 1 - y$

$\boxed{y = \frac{4}{7}}$ $\qquad$ $\boxed{x = 3/7}$

$$\therefore v(x,y) = \left[\underset{W}{3/7} \quad \underset{L}{4/7}\right]$$

In the long run probability of winning is $3/7$

→ Consider $E(x^2) = \Sigma x^2 . f(x)$

$= 0^2(1/2) + 1^2(1/2) = 1/2$

① becomes,

$\sigma_x^2 = 1/2 - (1/2)^2$

$= 1/2 - 1/4$

$= \boxed{1/4}$

→ Consider $E(Y)^2 = \Sigma Y^2 . f(Y)$

$= 0^2(1/8) + 1^2(3/8) + 2^2(3/8) + 3^2(1/8)$

$= 3$

② becomes,

$\sigma_{Y^2} = 3 - (3/2)^2 = 3 - 9/4 = \frac{12-9}{4} = \boxed{3/4}$

Co-variance :-

$\mathrm{COV}(XY) = E(XY) - E(X) . E(Y)$

$= 1/2 - 1/2(3/2)$

$= \boxed{-1/4}$

Co-relation :-

$$\rho(XY) = \frac{\mathrm{cov}(XY)}{\sigma_x \sigma_Y}$$

$$= \frac{-1/4}{1/2(\sqrt{3}/2)} = \boxed{\frac{-1}{\sqrt{3}}}$$

MODULE 5 : ANOVA

Anova test is used to compare more than two population (or) those population which are having more than two sample.

Assumption :- → Each population having normal distribution

→ Each sample should be independent & random

→ The population from which the sample are drawn have equal variances.

One Way Anova :- [Only 1 Factor]

Working Procedure :-

① Grand Total

$T = \Sigma X_1 + \Sigma X_2 + \Sigma X_3$

② Correction Factor

$C = \frac{T^2}{N}$ where T = total sum of all observations

N = total no. of observations

③ SSC [Sum of the squares b/w the samples]

$$SSC = \frac{(\Sigma X_1)^2}{n} + \frac{(\Sigma X_2)^2}{n} + \frac{(\Sigma X_3)^2}{n} - \frac{T^2}{N}$$

④ SST [Total sum of squares]

$$SST = \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^3 - \frac{T^2}{N}$$

⑤ SSE [Sum of the squares within the samples]

$SSE = SST - SSC$

Square of variation	Sum of squares	Degrees of freedom	Mean squares	F-ratio
B/w the samples	SSC	C-1	$MSC = \frac{SSC}{C-1}$	$F = \frac{MSC}{MSE}$
Within the samples	SSE	N-C	$MSE = \frac{SSE}{N-C}$	
Total	SST	N-1		

① It is desired to compare three hospitals with regards to the no. of deaths per quarter. A sample of death records were selected from the records of each hospitals & the no. of deaths was as given below. From these data, suggest a difference in the number of death per quarter among three hospitals. Given $F_{0.05}(2,12) = 3.89$.

Hospital 1	Hospital 2	Hospital 3
8	7	12
10	5	9
7	10	13
14	9	12
11	9	4

$\Sigma x_1 = 50$ $\quad \Sigma x_2 = 40$ $\quad \Sigma x_3 = 60$

→ Hypothesis : There is no significant difference in the no. of deaths among 3 hospitals

i.e., $\mu_1 = \mu_2 = \mu_3$

① Grand Total

$$T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3$$

$$= 50 + 40 + 60$$

$$= \boxed{150}$$

② Correction Factor

$$\frac{T^2}{N} = \frac{150 \times 150}{15} = \boxed{1500}$$

③ SSC

$$SSC = \frac{(\Sigma x_1)^2}{n} + \frac{(\Sigma x_2)^2}{n} + \frac{(\Sigma x_3)^2}{n} - \frac{T^2}{N}$$

$$= \frac{(50)^2}{5} + \frac{(40)^2}{5} + \frac{(60)^2}{5} - 1500$$

$$\boxed{SSC = 40}$$

④ SST

$$SST = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 - \frac{T^2}{N}$$

$$= 530 + 336 + 734 - (1500)$$

$$\boxed{SST = 100}$$

⑤ SSE

$$SSE = SST - SSC$$

$$= 100 - 40$$

$$\boxed{SSE = 60}$$

Source of variation	Sum of squares	Degrees of freedom		
B/w the samples	SSC = 40	$C = 3-1 = 2$	$MSC = \frac{SSC}{C-1} = 20$	$F = \frac{MSC}{MSE} = \frac{20}{5} = 4$
Within the samples	SSE = 60	$N-C = 15-3 = 12$	$MSE = \frac{SSE}{N-C} = 5$	
Total	SST = 100	$N-1 = 15-1 = 14$		

$\therefore F = 4 > F_{0.05}(2, 12) = 3.89$

Hence 95% confidence that our hypothesis is rejected i.e., 5% level of significance

② A manufacturing Company has purchased 3 new machines (A,B,C) of different makes & wishes to determine whether one of them is faster than the other in producing a certain item. From hourly production figures are observed at random from each machine & results are given below.

A	B	C
20	18	25
21	20	28
23	17	22
16	25	28
20	15	32

Use analysis of variance to test whether machine differ significantly. [Table value of F at 5% level for $V_1 = 2$ & $V_2 = 12$ is 3.89]

→ Hypothesis: $\mu_1 = \mu_2 = \mu_3$

$\Sigma x_1 = 100$, $\Sigma x_2 = 95$, $\Sigma x_3 = 135$

① Grand Total

$T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3$

$= 100 + 95 + 135$

$= \boxed{330}$

② Correction Factor

$\frac{T^2}{N} = \frac{(330)^2}{15} = \boxed{7260}$

③ SSC

$SSC = \frac{(\Sigma x_1)^2}{n} + \frac{(\Sigma x_2)^2}{n} + \frac{(\Sigma x_3)^2}{n} - \frac{T^2}{N}$

$= \frac{(100)^2}{5} + \frac{(95)^2}{5} + \frac{(135)^2}{5} - 7260$

$= \boxed{190}$

④ SST

$SST = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 - \frac{T^2}{N}$

$\boxed{SST = 330}$

⑤ SSE

$SSE = SST - SSC$

$= 330 - 190$

$\boxed{SSE = 140}$

Source of variation	Sum of squares	Degrees of freedom		
B/w the samples	$SSC = 190$	$C-1 = 3-1 = 2$	$MSC = \frac{SSC}{C-1} = \frac{190}{2} = 95$	$\frac{MSC}{MSE} = \frac{95}{11.67} = 8.14$
within the samples	$SSE = 140$	$N-C = 15-3 = 12$	$MSE = \frac{SSE}{N-C} = 11.67$	
Total	$SST = 330$	$N-1 = 15-1 = 14$		

$\therefore F = 8.14 > F_{0.05}(2.12) = 3.89$

Hence 95% confidence that our hypothesis is rejected i.e., 5% level of significance

Two Way Anova :- [2 Factors]

Working Procedure :-

① Grand Total

$$T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3$$

② Correction factor

$$\frac{T^2}{N}$$

③ SSC

$$SSC = \frac{(\Sigma x_1)^2}{n} + \frac{(\Sigma x_2)^2}{n} + \frac{(\Sigma x_3)^2}{n} - \frac{T^2}{N}$$

④ SSR

$$SSR = \frac{(\Sigma x_4)^2}{n} + \frac{(\Sigma x_5)^2}{n} + \frac{(\Sigma x_6)^2}{n} - \frac{T^2}{N}$$

⑤ SST

$$SST = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 - \frac{T^2}{N}$$

⑥ SSE

$$SSE = SST - [SSR + SSC]$$

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F-ratio
B/w the columns	SSC	$c-1$	$MSC = \frac{SSC}{c-1}$	$F_C = \frac{MSC}{MSE}$
B/w the rows	SSR	$r-1$	$MSR = \frac{SSR}{r-1}$	
Residuals	SSE	$(c-1)(r-1)$	$MSE = \frac{SSE}{(c-1)(r-1)}$	$F_R = \frac{MSR}{MSE}$

12M
***

① To study the performance of three detergents & three water temperatures the following whitness reading were obtained with specially designed equipment.

Water temperature	Detergent A	Detergent B	Detergent C
Cold water	57	55	67
Warm water	49	52	68
Hot water	54	46	58

Perform a two way analysis using 5% level of significance [Given F at 5% = 6.94]

→ Hypothesis :- 1. There is no difference in b/w whitenes due to three varieties of detergents

2. There is no diff in whiteness due to three temperatures

$\Sigma x_1 = 160$ $\Sigma x_2 = 153$ $\Sigma x_3 = 193$

$\Sigma x_4 = 179$ $\Sigma x_5 = 169$ $\Sigma x_6 = 158$

① Grand Total

$T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3$

$= 160 + 153 + 193$

$= \boxed{506}$

② Correction factor

$$\frac{T^2}{N} = \frac{(506)^2}{9} = 28\,448.44$$

• $SSC = \frac{(\Sigma x_1)^2}{n} + \frac{(\Sigma x_2)^2}{n} + \frac{(\Sigma x_3)^2}{n} - \frac{T^2}{N}$

$= \frac{(160)^2}{3} + \frac{(153)^2}{3} + \frac{(193)^2}{3} - 28\,448.44$

$\boxed{SSC = 304.22}$

• $SSR = \frac{(\Sigma x_4)^2}{n} + \frac{(\Sigma x_5)^2}{n} + \frac{(\Sigma x_6)^2}{n} - \frac{T^2}{N}$

$= \frac{(179)^2}{3} + \frac{(169)^2}{3} + \frac{(158)^2}{3} - 28\,448.44$

$\boxed{SSR = 73.56}$

• SST

$SST = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 - \frac{T^2}{N}$

$= 8566 + 7845 + 12477 - [28448.44]$

$= \boxed{439.56}$

• SSE

SSE = SST +-[SSR + SSC]

= 439.56 − [304.22 + 73.56]

$\boxed{SSE = 61.78}$

Source of variation	Sum of squares	Degrees of freedom	Mean Squares	F-ratio
B/w the columns	SSC	C-1 = 3-1 = 2	$MSC = \frac{SSC}{C-1} = 152.11$	$F_C = \frac{MSC}{MSE} = 9.85$
B/w the rows	SSR	r-1 = 3-1 = 2	$MSR = \frac{SSR}{r-1} = \frac{73.56}{2} = 36.78$	$F_R = \frac{MSR}{MSE} = 2.38$
Residual	SSE	(C-1)(r-1) = 4	$MSE = \frac{SSE}{(C-1)(r-1)} = \frac{61.78}{4} = 15.44$	

Case 1 : $F_C = 9.85 > F_{0.05}(2.4) = 6.94$

95% confidence that our First hypothesis is rejected

i.e., there is a significant diff b/w the detergents

Case 2 : $F_R = 2.38 < F_{0.05}(2.4) = 6.94$

95% confidence that our second hypothesis is accepted

i.e., there is no significant diff b/w water temperature

② The following data represents the no. of units of a product by 3 different workers using 3 different types of machines.

Workers	Machine A x_1	Machine B x_2	Machine C x_3	
X	8	32	20	$\Sigma x_4 = 60$
Y	28	36	38	$\Sigma x_5 = 102$
Z	6	28	14	$\Sigma x_6 = 48$
	$\Sigma x_1 = 42$	$\Sigma x_2 = 96$	$\Sigma x_3 = 72$	

Test (i) whether the mean productivity is the same for the different machine types.

(ii) whether the three workers differ with respect to mean productivity

[Given $F_{0.05}(2.4) = 6.94$]

→ Grand Total

$T = \Sigma x_1 + \Sigma x_2 + \Sigma x_3$

$= 42 + 96 + 72$

$= \boxed{210}$

② Correction Factor

$$\frac{T^2}{N} = \frac{(210)^2}{9} = 4900$$

③ SSC

$$SSC = \frac{(\Sigma x_1)^2 + (\Sigma x_2)^2 + (\Sigma x_3)^2}{n} - \frac{T^2}{N}$$

$$= \frac{(42)^2 + (96)^2 + (72)^2}{3} - 4900$$

$\boxed{488}$

④ SSR :-

$$SSR = \frac{(\Sigma x_4)^2 + (\Sigma x_5)^2 + (\Sigma x_6)^2}{n} - \frac{T^2}{N}$$

$$= \frac{(60)^2 + (102)^2 + (48)^2}{3} - 4900$$

$$= \boxed{536}$$

⑤ SST :-

$$SST = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 - \frac{T^2}{N}$$

$$= 814 + 3104 + 2040 - 4900$$

$$= \boxed{1128}$$

⑥ SSE :-

$$SSE = SST - [SSC + SSR]$$

$$= 1128 - [488 + 536]$$

$$= \boxed{104}$$

Source of variation	Sum of the squares	Degrees of freedom	Mean squares	F-ratio
B/w columns	SSC = 488	C − 1 = 2	$MSC = \frac{SSC}{C-1} = 244$	$F_C = \frac{MSC}{MSE} = 9.38$
B/w the rows	SSR = 536	r − 1 = 2	$MSR = \frac{SSR}{r-1} = 268$	$F_R = \frac{MSR}{MSE} = 10.31$
Residual	SSE = 104	(C−1)(r−1) = 4	$MSE = \frac{SSE}{(C-1)(r-1)} = 26$	

$F_C = 9.38 > F_{(2,4)} = 6.94$ our hypothesis is rejected

$F_R = 10.31 > F_{(2,4)} = 6.94$ our hypothesis is rejected

Three way Annova [Latin Square]:-

① Analyse the variance in the following latin square of yields (in kgs) of paddy where A, B, C, D denote the different methods of cultivation.

D 122	A 121	C 123	B 122
B 124	C 123	A 122	D 125
A 120	B 119	D 120	C 121
C 122	D 123	B 121	A 122

Examine whether the different methods of cultivation have given significantly different yields. Given that $F_{3,6} = 4.76$.

→ Given : Four types of (column) land, Four types of seeds (row)
Four types of cultivation
↓
ABCD (letters)

Hypothesis :- ① There is no significant difference b/w the columns
② — " — " — " — the rows
③ — " — " — " — the letters

→ The no. 120 is near to all, ∴ subtract all by 120

D 2	A 1	C 3	B 2
B 4	C 3	A 2	D 5
A 0	B −1	D 0	C 1
C 2	D 3	B 1	A 2

squares the values

D 4	A 1	C 9	B 4
B 16	C 9	A 4	D 25
A 0	B 1	D 0	C 1
C 4	D 9	B 1	A 4

(1) Grand Total

$T = 8+6+6+10$

$= \boxed{30}$

(2) Correction Factor

$\frac{T^2}{N} = \frac{(30)^2}{16} = 56.25$

(3) SSC

$$SSC = \frac{(\Sigma x_1)^2 + (\Sigma x_2)^2 + (\Sigma x_3)^2 + (\Sigma x_4)^2}{n} - \frac{T^2}{N}$$

$$= \frac{(8)^2 + (6)^2 + (6)^2 + (10)^2}{4} - 56.25$$

$$= \boxed{2.75}$$

(4) SSR

$$SSR = \frac{(\Sigma x_5)^2 + (\Sigma x_6)^2 + (\Sigma x_7)^2 + (\Sigma x_8)^2}{n} - \frac{T^2}{N}$$

$$= \boxed{24.75}$$

(5) SSL

$$SSL = \frac{(\Sigma A)^2 + (\Sigma B)^2 + (\Sigma C)^2 + (\Sigma D)^2}{4} - \frac{T^2}{N}$$

$\Sigma A = 1+2+0+2 = 5$

$\Sigma B = 2+4+1+1 = 6$

$\Sigma C = 3+3+1+2 = 9$

$\Sigma D = 2+5+0+3 = 10$

$$= \frac{(5)^2 + (6)^2 + (9)^2 + (10)^2}{4} - 56.25$$

$$= \boxed{4.25}$$

(6) SST

$$SST = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 + \Sigma x_4^2 - \frac{T^2}{N}$$

$$= 92 - 56.25$$

$$= \boxed{35.75}$$

(7) SSE

$$SSE = SST - [SSC + SSR + SSL]$$

$$= 35.75 - [2.75 + 24.75 + 4.25]$$

$$= \boxed{4}$$

Source of variation	Sum of squares	Degree of freedom	Mean of sum of squares	F-ratio
Columns	$SSC = 2.75$	$C-1 = 4-1 = 3$	$\frac{SSC}{C-1} = 0.916$	$\frac{MSC}{MSE} = 1.36$
rows	$SSR = 24.75$	$= r-1 = 3$	$\frac{SSR}{r-1} = 8.25$	$\frac{MSR}{MSE} = 12.31$
letters	$SSL = 4.25$	$= l-1$, $4-1 = 3$	$\frac{SSL}{l-1} = 1.416$	$\frac{MSL}{MSE} = 2.113$
error	$SSE = 4$	$(4-1)(4-2)$, $3 \times 2 = 6$	$\frac{SSE}{6} = 0.67$	