PROBABILITY & STATISTICS

(EMA203)

1 A box contains 6 red, 4 white and 5 black balls. A person draws 4 balls from the box at random. Find the probability that among the balls drawn there is atleast one ball of each colour.

Solution:-

Given: A box contains,

no. of red balls = 6

no. of white balls = 4

no. of black balls = 5

no. of balls drawn at random = 4

Total no. of possible outcomes = $\boxed{{}^{15}C_4 = n(S)}$

Let A be the event of getting a ball of one colour each from the balls drawn,

$$n(A) = {}^6C_1 \cdot {}^4C_1 \cdot {}^5C_2 + {}^6C_1 \cdot {}^4C_2 \cdot {}^5C_1 + {}^6C_2 \cdot {}^4C_1 \cdot {}^5C_1$$

$$= 6 \cdot 4 \cdot 10 + 6 \cdot 6 \cdot 5 + 15 \cdot 4 \cdot 5$$

$$= 240 + 180 + 300$$

$$\boxed{n(A) = 720}$$

Then, $P(A) = \dfrac{n(A)}{n(S)}$

$$= \frac{720}{{}^{15}C_4}$$

$$= \frac{720}{1365}$$

$$\boxed{P(A) = 0.5274}$$

$\boxed{\therefore P(\text{getting a ball of atleast one colour}) = 0.5274}$

(2) If two dice are thrown, what is the probability that the sum is (1) greater than 8 (2) Atleast 8 (3) neither 7 nor 11

Solution:-

Given: Two dice are thrown,

No. of total possible outcomes = $\boxed{n(S) = 36}$

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

(1) Let A be event of getting a sum greater than 8

$A = \{ (6,3)(5,4)(6,4)(4,5)(5,5)(6,5)(3,6)(4,6)(5,6)(6,6) \}$

$n(A) = 10$

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

$\boxed{P(A) = \frac{10}{36} = \frac{5}{18}}$

(2) Let B be the event of getting a sum of atleast 8

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

$n(B) = 14$

$P(B) = \frac{n(B)}{n(S)}$

$\boxed{P(B) = \frac{15}{36} = \frac{5}{12}}$

(3) Let C be the event of getting sum of 7 or 11

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

$n(C) = 8$

$P(C) = \frac{n(C)}{n(S)}$

$P(C) = \frac{8}{36}$

$P(\text{neither } 7 \text{ nor } 11) = P(\bar{C}) = 1 - \frac{8}{36} = \frac{28}{36}$

$\boxed{P(\bar{C}) = \frac{7}{9}}$

3 The odds against that A speaks the truth are 3:2 and the odds that B speaks the 5:3. In what percentage of cases are they likely to contradict each other on an identical point?

Solution:-

Let X be an event of A speaks against the truth,

$$\boxed{P(X) = \frac{2}{5}}$$

Then, $P(\bar{X}) = \frac{3}{5}$

Let Y be an event of B that speaks the truth,

$$\boxed{P(Y) = \frac{5}{8}}$$

Then, $P(\bar{Y}) = \frac{3}{8}$

P(cases contradict each other on an identical path)

$= P[(X \cap Y^c) \cup (X^c \cap Y)]$

$= P(X \cap Y^c) + P(X^c \cap Y)$

$= P(X) \cdot P(Y^c) + P(X^c) P(Y)$ [$\because$ X, Y are independent events]

$= \frac{2}{5} \cdot \frac{3}{8} + \frac{3}{5} \cdot \frac{5}{8}$

$= \frac{21}{40}$

$$\boxed{P = 0.525}$$

$$\boxed{P(\text{cases contradict each other}) = 52.5\%}$$

4) A problem in statistics is given to three students A, B, C whose chances of solving it are respectively $\frac{1}{2}$, $\frac{1}{3}$ and $\frac{1}{4}$. What is the probability that the problem will be solved?

Solution:-

Given: $P(\text{student A solving a problem}) = \frac{1}{2} = P(A)$

$P(\text{student B solving a problem}) = \frac{1}{3} = P(B)$

$P(\text{student C solving a problem}) = \frac{1}{4} = P(C)$

Then,

$P(\text{solving the problem}) = P(A \cup B \cup C)$

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

$$= P(A) + P(B) + P(C) - P(A) \cdot P(B) - P(B)P(C) - P(C)P(A) + P(A) \cdot P(B) \cdot P(C)$$

[$\because$ A, B, C are independent events]

$$= \frac{1}{2} + \frac{1}{3} + \frac{1}{4} - \frac{1}{2} \cdot \frac{1}{3} - \frac{1}{3} \cdot \frac{1}{4} - \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{4}$$

$$= \frac{13}{12} - \frac{1}{6} - \frac{1}{12} - \frac{1}{8} + \frac{1}{24}.$$

$$= 1 - \frac{1}{6} - \frac{1}{8} + \frac{1}{24}$$

$$= \frac{25}{24} - \frac{14}{48}$$

$$= \frac{36}{48}$$

$$\boxed{P(A \cup B \cup C) = \frac{3}{4}}$$

5. In a bolt factory machines A_1, A_2, A_3 manufacture 25%, 35%, 40% of total output. Of these 5, 4 and 2 percent are defective bolts. A bolts is drawn at random from the product and is found to be defective. What is the probability that it was manufactured by A_2.

Solution:-

Let $P(A_1), P(A_2), P(A_3)$ be the probabilities of events that the bolts are manufactured by machines A_1, A_2, A_3 respectively. Then,

$$P(A_1) = \frac{25}{100} \; ; \; P(A_2) = \frac{35}{100} \; ; \; P(A_3) = \frac{40}{100}$$

Let X denote that the bolt is defective. Then,

$$P(X/A_1) = \frac{5}{100} = 0.05$$

$$P(X/A_2) = \frac{4}{100} = 0.04$$

$$P(X/A_3) = \frac{2}{100} = 0.02$$

If the bolt is defective, then the probability that it is from machine A_2.

$$\boxed{P(A_2/X) = \frac{P(X/A_2) \cdot P(A_2)}{P(X/A_1)P(A_1) + P(X/A_2)P(A_2) + P(X/A_3) \cdot P(A_3)}}$$

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

$$= \frac{35 \times 4}{25 \times 5 + 35 \times 4 + 40 \times 2}$$

$$= \frac{140}{125 + 140 + 80}$$

$$= \frac{140}{345}$$

$$\boxed{P(A_2/X) = 0.4057}$$

6. The chance that A will diagnose a disease x correctly is 60%. The chance that a patient will die by his treatment after correct diagnosis is 40%. and the chance of death by wrong diagnosis is 70%. A patient of doctor A, whose had disease x died. The chance that his disease was diagnosed correctly.

Solution:-

Let E_1 be the event that the disease x is diagnosed correctly by doctor A

E_2 be the event that disease x is diagnosed correctly but patient will die.

$$P(E_1) = \frac{60}{100} = 0.6$$

$$P(E_2/E_1) = \frac{40}{100} = 0.4$$

$$P(\bar{E}_1) = 1 - 0.6 = 0.4$$

$$P(E_2/\bar{E}_1) = \frac{70}{100} = 0.7$$

The Baye's theorem states that,

$$P(E_1/E_2) = \frac{P(E_1)P(E_2/E_1)}{P(E_2/E_1)P(E_1) + P(E_2/\bar{E}_1)P(\bar{E}_1)}$$

$$= \frac{0.6 \times 0.4}{0.6 \times 0.4 + 0.4 \times 0.7}$$

$$= \frac{24}{24 + 28}$$

$$= \frac{24}{52}$$

$$= \frac{6}{13}$$

$$\boxed{P(E_1|E_2) = \frac{6}{13}}$$

7 A random variable has the following probability distribution.

X	0	1	2	3	4	5	6	7	8	9
P(X)	a	3a	5a	7a	9a	11a	13a	15a	17a	19a

(1) Determine 'a'
(2) Find (i) $P(X<3)$ (ii) $P(X \leq 3)$ (iii) $P(X>7)$ (iv) $P(2 \leq X \leq 5)$
(v) mean and variance.

Solution:-

(1) Since $\sum_{i=0}^{9} P(x_i) = 1$, we have

$$\Rightarrow a+3a+5a+7a+11a+9a+13a+15a+17a+19a = 1$$

$$\Rightarrow 100a = 1$$

$$\Rightarrow \boxed{a = 0.01}$$

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

$= a+3a+5a$

$= 9(0.01)$

$$\boxed{P(X<3) = 0.09}$$

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

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

$= 16(0.01)$

$$\boxed{P(X \leq 3) = 0.16}$$

(iii) $P(X>7) = P(X=8)+P(X=9)$

$= 17a+19a$

$= 36(0.01)$

$$\boxed{P(X>7) = 0.36}$$

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

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

$= 32(0.01)$

$$\boxed{P(2 \leq X \leq 5) = 0.32}$$

(v) Mean $= \mu = \sum p_i x_i$

$\Rightarrow 0.03 + 0.10 + 0.21 + 0.36 + 0.55 + 0.78 + 1.05 + 1.36 + 1.71$

$\Rightarrow 6.15$

variance $= \sum p_i x_i^2 - \mu^2$

$\sum p_i x_i^2 = 0.03 + 0.2 + 0.63 + 1.44 + 2.75 + 4.68 + 7.35 + 10.88 + 15.39$

$= 43.35$

$\sigma^2 = \sum p_i x_i^2 - \mu^2$

$= 43.35 - 37.8225$

$= 5.5275$

$\boxed{\sigma^2 = 5.5275}$

8. A random variable X has the following probability distribution

X	-3	6	9
P(X=x)	1/6	1/2	1/3

(1) Mean and variance

(2) Find E(Y), var(Y) given that $Y = 2X+1$

Solution:-

(1) Mean $= \mu = \sum x P(X=x)$

$= -3\left(\frac{1}{6}\right) + 6\left(\frac{1}{2}\right) + \frac{1}{3}(9)$

$= -0.5 + 3 + 3$

$\boxed{\mu = 5.5}$

variance $= \sum x^2 P(X=x) - \mu^2$

$= \frac{9}{6} + \frac{36}{2} + \frac{81}{3} - 30.25$

$= 1.5 + 18 + 27 - 30.25$

$= 46.5 - 30.25$

$\boxed{\sigma^2 = 16.25}$

(2) $E(Y) = E(2X+1)$

$= 2E(X)+1$

$= 2(5.5)+1$

$= 11.0+1$

$= 12$

$E(Y) = 12$

$var(Y) = 4E(X)$

$= 4(16.25)$

$= 65$

$\boxed{var(Y) = 65}$

4 A random variable X has the following probability function.

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

Find (i) K (ii) $P(X<6)$ (iii) $P(X\geq 6)$ (iv) c, $P(X\leq c)>1/2$ (v) Mean, S.D

Solution:-

(i) Since $\sum_{i=0}^{7} P(X=x) = 1$, we have

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

$10K^2+9K = 1$

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

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

$\boxed{K = \frac{1}{10}}$ $[\because 0 \leq P(x) \leq 1]$

(ii) $P(X<6) = 1 - P(X=7)$

$= 1-(7K^2+K)$

$= 1-(7(0.01)+0.1)$

$= 1-0.17$

$= 0.83$

$\boxed{P(X<6) = 0.83}$

(iii) $P(X>6) = P(X=7)$

$= 7K^2+K$

$= 7(0.01)+0.1$

$\boxed{P(X>6) = 0.17}$

(iv) $P(X\leq c) > 1/2$

$F(0) = P(X\leq 0) = P(X=0) = 0$

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

$F(2) = P(X\leq 2) = P(X=0)+P(X=1)+P(X=2) = \frac{3}{10}$

$F(3) = P(X\leq 3) = P(X=0)+P(X=1)+P(X=2)+P(X=3) = \frac{5}{10}$

$F(4) = P(X\leq 4) = P(X=0)+P(X=1)+P(X=2)+P(X=4) = \frac{8}{10}$

$F(5) = P(X\leq 5) = P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)+P(X=5) = \frac{81}{100}$

$F(6) = P(X\leq 6) = 1-P(X=7) = \frac{83}{100}$

$F(7) = P(X\leq 7) = 1$

The value of c for which $P(X\leq c) > \frac{1}{2}$ is minimum

$\boxed{K = 4, 5, 6, 7}$

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

$= 0+K+4K+6K+12K+5K^2+12K^2+49K^2+7K$

$= 66K^2+30K$

$= 0.66+3$

$\boxed{\mu = 3.66}$

variance $= \sum_{i=0}^{7} p_i x_i^2 - \mu^2$

$= K+8K+18K+48K+25K^2+72K^2+343K^2+49K-(3.66)^2$

$= 440K^2+124K-(3.66)^2$

$\boxed{\text{variance} = 3.4044}$

Generated by CamScanner

10) A continous random variable X has the probability density law $f(x) = Ax^2, 0 \le x \le 1$. Determine (i) A (ii) $P(X \le a) = P(X > a)$ (iii) Mean, S.D

Solution:-

Given: $f(x)$ is a continous Distribution Function since X is a continous Random variable.

(i) Since the total probability is unity,

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

$$\int_0^1 A(x^2)\,dx = 1$$

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

$$\boxed{A = 3}$$

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

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

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

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

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

$$\boxed{a^3 = \frac{1}{\sqrt[3]{2}}}$$

(iii) Mean, $\mu = \int_0^1 x f(x)\,dx = \int_0^1 3x^3 = 3\left[\frac{x^4}{4}\right]_0^1 = \frac{3}{4}$

$$\boxed{\mu = 0.75}$$

iv variance $= E(X^2) - [E(X)]^2$

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

$$= 3\left[\frac{x^5}{5}\right]_0^1 - 0.5625$$

$$= 0.6 - 0.5625$$

$$\sigma^2 = 0.0375$$

$$S.D = \sigma = \sqrt{\sigma^2} = \sqrt{0.0375}$$

$$\boxed{S.D = 0.1936}$$

(11) The diameter X of an electric cable is assumed to be continuous random variable with pdf $f(x) = 6x(1-x)$, $0 \le x \le 1$

(a) Show that it is PDf

(b) determine b such that $P(X \le b) = P(X > b)$

(c) Mean, S.D

Solution:-

Given : X is a continous random variable with probability density function $f(x) = 6x(1-x)$.

(a) For an equation to be PDf it must satisfy the following conditions,

(i) $f(x) > 0 \ \forall x \in R$

$6x(1-x) > 0$

$x(1-x) > 0$

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

$= \int_0^1 6x(1-x)\,dx$

$= 6\int_0^1 x - x^2\,dx$

$= 6\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1$

$= \cancel{6} \cdot \frac{1}{\cancel{6}}$

$= 1$ $\boxed{\therefore f(x) \text{ is PDf}}$

(b) $P(X \le b) = P(X > b)$

$\int_0^b \cancel{6}x(1-x) = \int_b^1 \cancel{6}x(1-x)$

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

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

$6b^2 - 4b^3 = 1$

$b^2(6-4b) = 1$

$\boxed{b = \pm 1, \frac{3}{2}}$

(c) Mean, E(X)

$\int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 6x^2(1-x)\,dx$

$= 6\int_0^1 x^2 - x^3\,dx$

$$= 6\left[\frac{x^3}{3}-\frac{x^4}{4}\right]_0^1$$

$$= 6\left[\frac{1}{12}\right]$$

$$\boxed{E(X) = \frac{1}{2} = 0.5}$$

$$E(X^2) = \int_0^1 x^2 f(x)\,dx = \int_0^1 6x^3(1-x)\,dx$$

$$= 6\left[\frac{x^4}{4}-\frac{x^5}{5}\right]_0^1$$

$$= 6\left[\frac{1}{4}-\frac{1}{5}\right]$$

$$= \frac{3}{10}$$

$$\text{var}(X) = \frac{3}{10} - \left(\frac{1}{2}\right)^2 = \frac{1}{20}$$

$$\boxed{S.D = \frac{1}{\sqrt{20}} = 0.2236}$$

12 A continous R.V X has the distribution $F(X) = \begin{cases} 0 & x \le 1 \\ K(x-1)^4 & 1 < x \le 3 \\ 1 & x > 3 \end{cases}$

(i) pdf

(ii) Determine K

(iii) Mean & S.D

Solution:-

$$f(x) = \frac{d}{dx}[F(x)]$$

$$f(x) = \begin{cases} 0 & x \le 1 \\ 4K(x-1)^3 & 1 < x \le 3 \\ 0 & x > 3 \end{cases}$$

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

$$\int_1^3 4K(x-1)^3\,dx = 1$$

$$4K\left[\frac{(x-1)^4}{4}\right]_1^3 = 1$$

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

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

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

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

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

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

Let $x-1=t$

$= \frac{1}{4} \int_{0}^{2} (t+1) t^3\, dx$

$= \frac{1}{4} \left(\frac{t^5}{5} + \frac{t^4}{4} \right)_0^2$

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

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

$= 4 \left[\frac{13}{30} \right]$

$\boxed{E(X) = 2.6}$

variance $[E(X^2)] = \int_{0}^{3} x^2 f(x) - \left[\frac{26}{15} \right]^2$

Let $x-1=t$, $= \int_{0}^{2} (t+1)^2 t^3 dt - 6.76$

$= \left[\frac{t^6}{6} + \frac{2t^5}{5} + \frac{t^4}{4} \right]_0^2 - 6.76$

$= \frac{64}{6} + \frac{64}{5} + \frac{16}{4} - 6.76$

$= \frac{1}{4}(27.4) - 6.76$

$= 0.09$

S.D $(\sigma) = \sqrt{0.09}$

$\boxed{\sigma = 0.3}$

[P3] A multiple choice questionnaire has 12 questions with 5 options A-E. A student is completely unprepared and wrote the test. If test follows binomial probability law find

(1) Exactly two answers are correct

(2) Atleast two answers are correct

(3) More than 7 answers are correct

Solutions-

By Binomial distribution, we know that

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

p = probability of getting an answer $= \frac{1}{5}$

q = probability of not getting an answer $= \frac{4}{5}$

$n = 12$

(i) $P(X=2) = {}^{12}C_2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^{12-2}$

$= {}^{12}C_2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^{10}$

$= 66 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^{10}$

$= 0.1073 \times 66 \times 0.04$

$= 0.2834$

$$\boxed{P(X=2) = 0.2834}$$

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

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

$= 1 - \left[{}^{12}C_0 \left(\frac{1}{5}\right)^0 \left(\frac{4}{5}\right)^{12} + {}^{12}C_1 \left(\frac{1}{5}\right)^1 \left(\frac{4}{5}\right)^{11}\right]$

$= 1 - [0.0684 + 0.2061]$

$= 1 - 0.274$

$= 0.725$

$$\boxed{P(X \geq 2) = 0.725}$$

(iii) $P(X>7) = 495\left(\frac{1}{5}\right)^8\left(\frac{4}{5}\right)^4 + 220\left(\frac{1}{5}\right)^7\left(\frac{4}{5}\right)^5 + 66\left(\frac{1}{5}\right)^{10}\left(\frac{4}{5}\right)^2$

$+ 12\left(\frac{1}{5}\right)^{11}\left(\frac{4}{5}\right)^1 + 1.\left(\frac{1}{5}\right)^{12}\left(\frac{4}{5}\right)^0$

$= 5.190\times10^{-4} + 5.676\times10^{-5} + 4.325\times10^{-6} + 1.966\times10^{-7} + 4.096\times10^{-9}$

$= 5.8119\times10^{-4}$

$$\boxed{P(X>7) = 0.00058119 = 0.05\times10^{-2}}$$

(14) With the usual notation find p for binomial R.v X if $n=6$ and $9P(X=4)=P(X=2)$

Solution:-

By Binomial Distribution, we know that

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

$$9P(X=4) = P(X=2)$$

$$9\,{}^6C_4\, p^4q^2 = {}^6C_2\, p^2q^4$$

$$\frac{p^{\cancel{4}\,2}\cancel{q^2}}{\cancel{p^2}\,q^{\cancel{4}\,2}} = \frac{1}{9}$$

$$\left(\frac{p}{q}\right)^2 = \frac{1}{9}$$

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

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

$$3p = 1-p \quad [\because q = 1-p]$$

$$4p = 1$$

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

15. The mean and variance of a binomial variate X with parameters n and p are 16 and 8. Find $P(X=1)$.

Solutions:-

Given: mean = 16

variance = 8

$np = 16$ — ①

$npq = 8$ — ②

Dividing ①/②

$$\frac{\cancel{np}}{\cancel{np}q} = \frac{16^{2}}{8}$$

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

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

$$n\left(\frac{1}{2}\right) = 16$$

$$n = 32$$

$$P(X=1) = {}^{32}C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^{31}$$

$$= 32 \cdot \left(\frac{1}{2}\right)^{32}$$

$$\boxed{P(X=1) = 7.4505 \times 10^{-9}}$$

16. A typist makes an average 2 mistakes per page. what is probability of a particular page having no errors in it?

Solution:-

By possion distribution,

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

Here, $\lambda = 2$

$$\therefore P(X=0) = \frac{e^{-2} 2^0}{0!}$$

$$= 0.13533$$

$$\boxed{P(X=0) = 0.13533.}$$

17) A textbook has 390 typographical errors of 520 pages. If it follows poisson law. Find the probablity that
(i) A page contains exactly one error.
(ii) A page contains exactly two errors,
(iii) A page contains more than 3 errors.
Also find the probability that 5 pages contain no error. Find the expected no. of pages which contains at least two errors.

solution :-

Let $X \sim P(\lambda)$. Then, $\boxed{P(X=x) = \frac{e^{-\lambda} \lambda^x}{x!}}$

$$\text{mean} = \lambda = \frac{390}{520} = 0.75$$

(i) $P(X=0) = \frac{e^{-0.75} \cdot (0.75)^0}{0!}$

$$\boxed{P(X=0) = 0.47}$$

(ii) $P(X=1) = \frac{e^{-0.75} (0.75)^1}{1!}$

$$\boxed{P(X=1) = 0.35}$$

(iii) $P(X \geq 1) = 1 - P(x<1)$

$= 1 - P(X=0)$

$= 1 - 0.47$

$= 0.53$

(iv) $P(X=2) = \dfrac{e^{-0.75}(0.75)^2}{2!} = 0.132$

$\boxed{P(X=2) = 0.132}$

(v) P(5 pages contains no errors) $= [P(X=0)]^5$

$= (0.47)^5$

$= 0.02293$

(vi) Expected frequency $= N \times P(X=x)$

$= 520 \times P(X=2)$

$= 520 \times 0.132$

$\boxed{\text{Expected frequency} = 68.9}$

18. Components are packed in boxes of 20. The probability of a component being defective is 0.1. What is the probability of a box containing 2 defective components.

Solution:-

Given: no. of boxes = 20

Probability of a component being defective $= p = 0.1$

Then, mean = np

$\lambda = 20 \times 0.1 = 2$

By poissoin law,

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

Then, the probability of a box containing 2 errors,

$P(X=2) = \dfrac{e^{-2}\,2^2}{2!}$

$= 0.135 \times 2$

$= 0.2706$

$\boxed{P(X=2) = 0.2706}$

17) If a R.V X follows P.D with parameter λ given that $P(X=2)=9P(X=4)+90P(X=6)$. Then, find the parameter

(i) λ

(ii) $P(X\geq 1)$

(iii) $P(0\leq X\leq 2)$

Solution:-

given : $P(X=2)=9P(X=4)+90P(X=6)$

(i) $P(X\geq 1)$

By poisson distribution,

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

$$P(X=2)=\frac{e^{-\lambda}\lambda^{2}}{2!}$$

$$P(X=4)=\frac{e^{-\lambda}\lambda^{4}}{4!}$$

$$P(X=6)=\frac{e^{-\lambda}\lambda^{6}}{6!}$$

Then, $P(X=2)=9P(X=4)+90P(X=6)$

$$\frac{e^{-\lambda}\lambda^{2}}{2!}=9\frac{e^{-\lambda}\lambda^{4}}{4!}+90\frac{e^{-\lambda}\lambda^{6}}{6!}$$

$$\frac{e^{-\lambda}\lambda^{2}}{2!}\left[\frac{90\lambda^{4}}{360}+\frac{9\lambda^{2}}{12}-1\right]=0$$

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

$$e^{-\lambda}\lambda^{2}(\lambda^{2}+4)(\lambda^{2}-1)=0$$

$$\lambda^{2}-1=0\ [e^{-\lambda}>0,\ \lambda>0\ \&\ \lambda^{2}-4\neq 0]$$

$$\boxed{\lambda=1\ (\lambda>0,\ \lambda\neq 1)}$$

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

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

$$=1-0.3678$$

$$\boxed{P(X\geq 1)=0.6321}$$

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

$$=\frac{e^{-1}1^{0}}{0!}+\frac{e^{-1}1^{1}}{1!}+\frac{e^{-\lambda}1^{2}}{2!}=\boxed{0.9196}$$

(20) X is normally distributed with mean $\mu = 30$ and S.D = 4
Find (a) $P(X<40)$ (b) $P(X>21)$ (c) $P(30<x<35)$

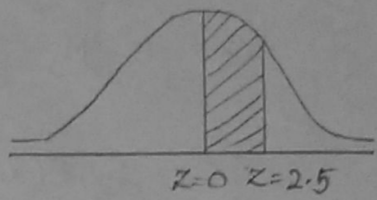
Solution:-

As it is normally distributed,

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

Given: $\mu = 30$, $\sigma = 4$

(a) when $x = 40$



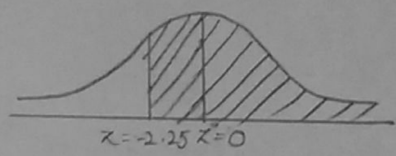
$$z = \frac{40-30}{4} = \frac{10}{4} = 2.5$$

$$P(X<40) = P(z<2.5)$$

$$\boxed{P(X<40) = 0.4938}$$

(b) when $x = 21$

$$z = \frac{21-30}{4} = -2.25$$



$$P(X>21) = P(z>-2.25)$$
$$= 0.5 + P(z<-2.25)$$
$$= 0.5 + 0.4878$$

$$\boxed{P(X>21) = 0.9822}$$

(c) $P(30<X<35)$

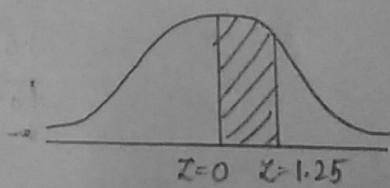
when $x = 30$

$$z = \frac{30-30}{4} = 0$$

when $x = 35$

$$z = \frac{35-30}{4} = 1.25$$

$$P(30<X<35) = P(0<z<1.25)$$



$$\boxed{P(30<X<35) = 0.3944}$$

(21) In normal distribution 10.03% of items are under 25kg and 89.97% of items are under 70kg. What are mean and S.D of the distribution?

Given : $P(x<25) = 10.03\% = 0.1003$

$P(x<70) = 89.97\% = 0.8997$

when $x = 25$,

$$z_1 = \frac{25-\mu}{\sigma} \quad \text{--- (1)}$$

when $x = 70$,

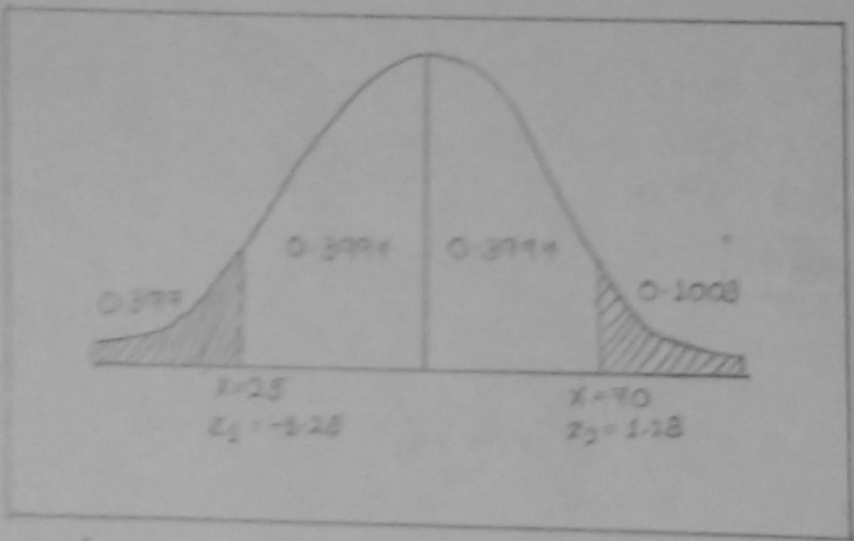
$$z_2 = \frac{70-\mu}{\sigma} \quad \text{--- (2)}$$

$P(x \le 70) = 1 - P(x<70) = 0.1003$

$P(z<z_2) = 0.3997 \Rightarrow z_2 = 1.28$

$P(x \le 25) = 1 - P(x<25) = 0.8997$

$P(z<z_1) = 0.3997 \Rightarrow z_1 = -1.28$



Substituting (1) and (2)

$$\begin{array}{r} 25 - \cancel{\mu} = -1.28\sigma \\ -70 \pm \cancel{\mu} = {}^{-}1.28\sigma \\ \hline \cancel{-}45 = \cancel{-}2.56\sigma \end{array}$$

$$\boxed{\sigma = 17.578 \text{ kgs}}$$

$$25 - \mu = -1.28(17.578)$$

$$\boxed{\mu = 47.499 \text{ kgs}}$$

23) Using least square method fit a curve of the form $y = ax^2 + bx + c$

x	1	3	5	6
y	12	18	25	35

Solution:-

Let $\boxed{y = ax^2 + bx + c}$ — (1) be the equation of the parabola.

The normal equations are

$$\sum_{i=1}^{n} y_i = nc + b\sum_{i=1}^{n} x_i + a\sum_{i=1}^{n} x_i^2 \quad \text{— (2)}$$

$$\sum_{i=1}^{n} x_i y_i = c\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 + a\sum_{i=1}^{n} x_i^3 \quad \text{— (3)}$$

$$\sum_{i=1}^{n} x_i^2 y_i = c\sum_{i=1}^{n} x_i^2 + b\sum x_i^3 + a\sum_{i=1}^{n} x_i^4 \quad \text{— (4)}$$

x	y	x^2	x^3	x^4	xy	x^2y
1	12	1	1	1	12	12
3	18	9	27	81	54	162
5	25	25	125	625	125	625
6	35	36	216	1296	210	1260
$\Sigma x = 15$	$\Sigma y = 90$	$\Sigma x^2 = 71$	$\Sigma x^3 = 369$	$\Sigma x^4 = 2003$	$\Sigma xy = 401$	$\Sigma x^2y = 2059$

$90 = 4c + 15b + 71a$ — (5)

$401 = 15c + 71b + 369a$ — (6)

$2059 = 71c + 369b + 2003a$ — (7)

By solving (5), (6) & (7),

$c = 12.298$

$b = -0.639$

$a = 0.7079$

$\boxed{\therefore y = 0.7079x^2 - 0.639x + 12.298}$

Q) Fit an exponential curve $y = ab^x$ for the following data.

X	40	65	90	5	30	10	80	85	70	25
Y	30	20	10	80	40	65	15	15	20	50

Solution:-

Let $\boxed{y = ab^x}$ — ① be a power function.

Then, applying log on both sides

$$\log y = \log a + x \log b$$

$$Y = A + Bx \quad \text{—} ②$$

$Y = \log y$; $A = \log a$; $B = \log b$

equation ② is a straight line.

The normal equations are

$$\sum_{i=1}^{n} Y = nA + B\sum_{i=1}^{n} x_i \quad \text{—} ③$$

$$\sum_{i=1}^{n} x_i Y_i = A\sum_{i=1}^{n} x_i + B\sum_{i=1}^{n} x_i^2 \quad \text{—} ④$$

x	y	$Y = \log y$	x^2	xY
40	30	1.4771	1600	59.08
65	20	1.3010	4225	84.565
90	10	1	8100	90
5	80	1.9030	25	9.515
30	40	1.6020	900	48.06
10	85	1.8129	100	18.129
80	15	1.1760	6400	94.08
85	15	1.1760	7225	99.96
70	20	1.3010	4900	91.07
25	50	1.6989	625	42.4725
$\sum x = 500$	$\sum y = 345$	$\sum Y = 14.43$	$\sum x^2 = 34100$	$\sum xY = 636.80$

$14.439 = 10A + 500B$ — ⑤

$636.805 = 500A + 34100B$ — ⑥

Solving ⑤ and ⑥

$A = 1.9247$

$B = -9.617 \times 10^{-3}$

$a = \text{antilog } A$

$a = 82.08$

$b = \text{antilog } B$

$b = 0.97809$

$\boxed{y = 82.08 \cdot (0.97809)^x}$ is the obtained power function.

Q24) Fit a power curve is the form of following data

x	1	2	3	4	5
y	8	15	30	60	125

Solution:-

Let $\boxed{y = ax^b}$ be the power curve

Applying log on both sides,

$\log y = \log a + b \log x$

$Y = A + bX$

$Y = \log y$; $A = \log A$; $\log x = X$

Thus, equation② is a straight line.

The normal equations are

$$\sum_{i=1}^{n} Y_i = nA + b\sum_{i=1}^{n} X_i \quad — ③$$

$$\sum_{i=1}^{n} X_iY_i = A\sum_{i=1}^{n} X_i + \sum_{i=1} X_i^2 \quad — ④$$

Generated by CamScanner

x	y	$X = \log x$	$Y = \log y$	XY	X^2
1	8	0	0.9030	0	0
2	15	0.3010	1.176	0.3539	0.09
3	30	0.4772	1.477	0.2045	0.227
4	60	0.6020	1.778	1.070	0.362
5	125	0.6989	2.096	1.464	0.4884
$\Sigma x = 15$	$\Sigma y = 238$	$\Sigma X = 2.078$	$\Sigma Y = 743$	$\Sigma XY = 3.59$	$\Sigma X^2 = 1.16$

Substituting the obtained values in ③ & ④,

$$7.43 = 5A + 2.0789b \quad -⑤$$

$$3.5932 = 2.0789A + 1.1679 \quad -⑥$$

$$A = 0.79$$

$$a = \text{antilog } A$$

$$a = 6.16$$

$$b = 1.66$$

$\therefore$ $\boxed{y = 6.16\,x^{1.66}}$ is the required power curve.

25) Calculate the correlation coefficient for the following heights of fathers and sons.

X	65	66	67	67	68	69	70	72
Y	67	68	65	68	72	72	69	71

Solution:-

correlation coefficient, $\boxed{r = \frac{\Sigma XY}{n \sigma_x \sigma_y}}$

$$\sigma_x = \sqrt{\frac{1}{n}\Sigma X^2} \quad ; \quad X = x - \bar{x}$$

$$\sigma_y = \sqrt{\frac{1}{n}\Sigma Y^2} \quad ; \quad Y = y - \bar{y}$$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{544}{8} = 68$$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{552}{8} = 69$$

x	y	$X = x - \bar{x}$	$Y = y - \bar{y}$	XY	X^2	Y^2
65	67	-3	-2	6	9	4
66	68	-2	-1	2	4	1
67	65	-1	-4	4	1	16
67	68	-1	-1	1	1	1
68	72	0	3	0	0	9
69	72	1	3	3	1	9
70	69	2	0	0	4	0
72	71	4	2	8	16	4
$\Sigma x = 544$	$\Sigma y = 552$	$\Sigma X = 0$	$\Sigma Y = 0$	$\Sigma XY = 24$	$\Sigma X^2 = 36$	$\Sigma Y^2 = 44$

Substituting the obtained values

$$r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \cdot \Sigma Y^2}}$$

$$r = \frac{24}{\sqrt{36 \times 44}}$$

$$= \frac{24}{39.79}$$

$\therefore$ $\boxed{r = 0.603}$ is the correlation coefficient

(26) Find the spearman rank correlation coeffienct to following data.

x	11	12	43	84	15
y	8	15	30	60	12

Solution:-

x	y	Rank of x	Rank of y	$d_i = x_i - y_i$	d_i^2
11	8	1	1	0	0
12	15	2	3	-1	1
43	30	4	4	0	0
84	60	5	5	0	0
15	12	3	2	1	1

$\Sigma d_i = 0 \quad \Sigma d_i^2 = 2$

The spearman's rank of correlation coefficient is given by

$$\rho = 1 - \frac{6 \Sigma d_i^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 2}{5(5^2-1)}$$

$$= 1 - \frac{12}{5 \times 24}$$

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

$$= \frac{9}{10}$$

$$\boxed{\rho = 0.9}$$

∴ Since, ρ lies between 0 and 1, its variables are partially positively correlated.

27) For the following data compute the correlation coefficient between x and y.

	x series	y series
No. of items :	15	15
Arthematic mean :	25	18
Sum of squares of derivations from mean :	136	138
Sum of products of derivations of x and y from mean :	122	

Solution:-

The correlation coefficient is given by

$$r(x,y) = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}}$$

$$= \frac{122}{\sqrt{136 \times 138}}$$

$$= \frac{122}{136.996}$$

$$= 0.8905$$

$$\therefore \boxed{r(x,y) = 0.8905}$$

28) Fit a linear regression equation of y on x to following data.

x	5	8	7	6	4
y	3	4	5	2	1

Solution:-

correlation coefficient, $\boxed{r = \frac{\Sigma XY}{n \sigma_x \sigma_y}}$

$\sigma_x = \sqrt{\frac{1}{n}\Sigma X^2} = \sqrt{\frac{1}{5} \times 10} = 1.414$; $X = x - \bar{x}$

$\sigma_y = \sqrt{\frac{1}{n}\Sigma Y^2} = \sqrt{\frac{1}{5} \times 10} = 1.414$; $Y = y - \bar{y}$

$\bar{x} = \frac{30}{5} = 6$

$\bar{y} = \frac{15}{5} = 3$

x	y	$X = x - \bar{x}$	X^2	$Y = y - \bar{y}$	Y^2	XY
5	3	-1	1	0	0	0
8	4	2	4	1	1	2
7	5	1	1	2	4	2
6	2	0	0	-1	1	0
4	1	-2	4	-2	4	4
$\Sigma x = 30$	$\Sigma y = 15$	$\Sigma X = 0$	$\Sigma X^2 = 10$	$\Sigma Y = 0$	$\Sigma Y^2 = 10$	$\Sigma XY = 8$

$$r = \frac{\Sigma xy}{n \sigma_x \sigma_y}$$

$$= \frac{8}{5 \times \sqrt{2} \times \sqrt{2}}$$

$$= \frac{8}{10}$$

$$r = 0.8$$

$$b_{yx} = r \cdot \frac{\sigma_y}{\sigma_x}$$

$$= 0.8 \times \frac{\sqrt{2}}{\sqrt{2}}$$

$$b_{yx} = 0.8$$

Line of regression y on x is given by

$$y - \bar{y} = b_{yx}(x - \bar{x})$$

$$y - 8 = 0.8(x - 6)$$

$$y = 0.8x - 4.8 + 3$$

$$\boxed{y = 0.8x - 1.8}$$

28) In a record of an analysis of correlation data, the following results are readable variance of $x = 9$, Regression equations. $8x - 10y + 66 = 0$ and $40x - 18y = 214$.

Find

(i) Mean values of x and y

(ii) correlation coefficient between x and y

(iii) the S.D of y

Solution :-

Given : $8x - 10y + 66 = 0$ — (1)

$40x - 18y - 214 = 0$ — (2)

(ii) Correlation coefficient,

$$r = \sqrt{b_{yx} \cdot b_{xy}}$$

Line of regression of y on x,

from ①, $-10y = -8x - 66$

$$y = \frac{8}{10}x + \frac{66}{10}$$

$$b_{yx} = \frac{8}{10} \quad — ③$$

Line of regression of x on y,

from ②, $40x = 18y + 214$

$$x = \frac{18}{40}y + \frac{214}{40}$$

$$b_{xy} = \frac{18}{40} \quad — ④$$

From ③ and ④ in r,

$$r = \pm\sqrt{\frac{8}{10} \times \frac{18}{40}}$$

$$\boxed{r = \pm\frac{3}{5} = 0.6}$$ [r is always positive]

(i) solving ① and ②,

$$\begin{array}{r} \cancel{40x} - 50y + 330 = 0 \\ -\ \cancel{40x} \pm 18y \pm 214 = 0 \\ \hline -32y + 544 = 0 \end{array}$$

$$y = 17$$

$$8x = 10(17) - 66$$

$$\boxed{x = 13.}$$

(iii) var of X = 9

$$b_{yx} = r\frac{\sigma_y}{\sigma_x}$$

$$0.8 = 0.6\,\frac{\sigma_y}{3}$$

$$\sigma_y = \frac{0.8 \times 3}{0.6} = 4$$

$$\boxed{\sigma_y = 4}$$

(30) Obtain the equations of two lines of regression for the data. Also obtain estimate of x for $y=70$

X	65	66	67	67	68	69	70	72
y	67	68	65	68	72	72	69	71

Solution:-

correlation coefficient, $r = \dfrac{\Sigma XY}{n\sigma_x \sigma_y}$

$\sigma_x = \sqrt{\frac{1}{n}\Sigma X^2}$; $X = x - \bar{x}$

$\sigma_y = \sqrt{\frac{1}{n}\Sigma Y^2}$; $Y = y - \bar{y}$

$\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{544}{8} = 68$

$\bar{y} = \dfrac{\Sigma y}{n} = \dfrac{552}{8} = 69$

x	y	$X = x-\bar{x}$	$Y = y-\bar{y}$	XY	X^2	Y^2
65	67	-3	-2	6	9	4
66	68	-2	-1	2	4	1
67	65	-1	4	-4	1	16
67	68	-1	-1	1	1	1
68	72	0	3	0	0	9
69	72	1	3	3	1	9
70	69	2	0	0	4	0
72	71	4	2	8	16	4
$\Sigma x = 544$	$\Sigma y = 552$	$\Sigma X = 0$	$\Sigma Y = 0$	$\Sigma XY = 24$	$\Sigma X^2 = 36$	$\Sigma Y^2 = 44$

$\sigma_x = \sqrt{\dfrac{36}{8}} = 2.12$

$\sigma_y = \sqrt{\dfrac{44}{8}} = 2.345$

$r = \dfrac{24}{8 \times 2.12 \times 2.345}$

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} = 0.603 \times \frac{2.345}{2.12}$$

$$= 0.6669$$

$$b_{xy} = r\frac{\sigma_x}{\sigma_y} = 0.603 \times \frac{2.12}{2.345}$$

$$= 0.545$$

Line of regression of y on x

$$y - \bar{y} = b_{yx}(x - \bar{x})$$

$$y - 69 = 0.6669(x - 68)$$

$$\boxed{y = 0.669x + 23.644}$$

when y = 70

$$70 = 0.669x + 23.644$$

$$x = 69.499$$

Line of regression x on y

$$x - \bar{x} = b_{xy}(y - \bar{y})$$

$$x - 68 = 0.545(y - 69)$$

$$\boxed{x = 0.545y - 30.395}$$

when y = 70

$$x = 0.545(70) - 30.395$$

$$x = 68.545$$

31) Estimate the production of 2010, by fitting regression line to following data

Year	2003	2004	2005	2006	2007
Production	5	8	14	12	13

Solution :-

correlation coefficient, $r = \frac{\sum XY}{n\sigma_x \sigma_y}$

$\sigma_x = \sqrt{\frac{1}{n}\sum X^2}$; $X = x - \bar{x}$

$\sigma_y = \sqrt{\frac{1}{n}\sum Y^2}$; $Y = y - \bar{y}$

$\bar{x} = \frac{\sum x}{n} = \frac{10025}{5} = 2005$

$\bar{y} = \frac{\sum y}{n} = \frac{52}{5} = 10.4$

Year	Production	$X = x - \bar{x}$	$Y = y - \bar{y}$	XY	X^2	Y^2
2003	5	-2	-5.4	10.8	4	29.16
2004	8	-1	-2.4	2.4	1	5.76
2005	14	0	3.6	0	0	12.96
2006	12	1	1.6	1.6	1	2.56
2007	13	2	2.6	5.2	4	6.76
$\sum x = 10025$	$\sum y = 52$	$\sum X = 0$	$\sum Y = 0$	$\sum XY = 20$	$\sum X^2 = 10$	$\sum Y^2 = 57.4$

$\sigma_x = \sqrt{\frac{1}{5} \times 10} = \sqrt{2}$

$\sigma_y = \sqrt{\frac{1}{5} \times 57.2} = 3.382$

$r = \frac{\sum xy}{n\sigma_x \sigma_y} = \frac{20}{5 \times 1.414 \times 3.382} = 0.836$

$b_{yx} = 0.836 \times \frac{3.382}{1.414} = 1.999$

Line of regression y on x,

$y - \bar{y} = b_{yx}(x - \bar{x})$

$y - 10.4 = 1.999(x - 2005)$

$\boxed{y = 1.999x - 3997.595}$

when $x = 2010$, then $y = 20.395$

$x = 2008$, then $y = 16.395$

$$b_{xy} = \frac{r\sigma_x}{\sigma_y} = \frac{0.836 \times 1.414}{3.382}$$

$$b_{xy} = 0.3495$$

Line of regression x on y,

$$x - \bar{x} = b_{xy}(y - \bar{y})$$

$$x - 2005 = 0.3495(y - 10.4)$$

$$\boxed{x = 0.3495y + 2001.3652}$$

when $x = 2010$, then $y = 24.706$

when $x = 2008$, then $y = 18.9836$.

32) A sample of 900 members has mean 3.4 cms and S.D 2.61 is the sample from a large population of mean 3.25 cms and S.D 2.61 cms? 5% level

Solution:-

Step1 - Setup for statistical hypothesis

μ = population mean

Null Hypothesis, $H_0 \Rightarrow \mu = 3.25$

Alternative hypothesis, $H_1 \Rightarrow \mu \neq 3.25$

∴ Therefore, it is a two-tailed test.

Step2 - Sample data

n = no. of sample members = 900

$\bar{x}$ = sample mean = 3.4

μ = population mean = 3.25

σ = standard deviation = 2.61

α = level of significance = 5% = 0.05

$Z_{tab} = 1.96$

Step3 - Test sample statistic

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

Step 4 - Calculations

$$z = \frac{3.4 - 3.25}{\frac{2.61}{\sqrt{900}}}$$

$$= \frac{0.15}{\frac{2.61}{30}}$$

$$\boxed{z = 1.724}$$

Step 5 - Conclusion

$z_{tab} = 1.96$

$z_{cal} = 1.724$

$\boxed{z_{cal} \leq z_{tab}}$

∴ Hence, the condition allows H_0 to be accepted.
Yes, the sample is from a large population of mean 3.25 cms and 2.61 standard deviation.

33 A sample of size 400 was drawn and the sample means B was found to be 99. Test whether this sample could have come from a normal population with mean 100 and variance 64 at 5% level of significance.

Solution :-

Step 1 - Setup for statistical hypothesis.

μ = population mean.

Null hypothesis, $H_0 \Rightarrow \mu = 100$

Alternative hypothesis, $H_1 \rightarrow \mu \neq 100$

∴ Therefore, it is a two-tailed test.

Step 2 - Sample data.

n = sample size = 400

$\bar{x}$ = sample mean = 99

μ = population mean = 100

σ = standard deviation = $\sqrt{64} = 8$

α = level of significance = 0.05

$z_{tab} = 1.96$

Step 3 - Test sample statistics

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

Step 4 - calculations

$$z = \frac{99 - 100}{\frac{8}{\sqrt{400}}}$$

$$= \frac{-1}{\frac{8}{20}}$$

$$\boxed{z = -2.5}$$

Step 5 - Conclusion

$|z_{cal}| = 2.5$

$z_{tab} = 1.96$

$$\boxed{z_{cal} > z_{tab}}$$

∴ Hence, the condition allows H_0 to be rejected. The sample could not have come from the normal population with mean 100.

34 The means of two large samples of 1000 and 2000 items are 67.5 cms and 68.0 cms respectively. Can the samples be regarded as drawn from from the population with S.D 2.5 cms. Test at 5% level of significance

Solution:-

Step 1 - Setup for statistical hypothesis.

μ_1 and μ_2 are means of two populations.

Null hypothesis, $H_0 \Rightarrow \mu_1 - \mu_2 = 0$

Alternative hypothesis, $H_1 \Rightarrow \mu_1 \neq \mu_2$

Step 2 - Sample data.

	sample I	sample II
sample size	$n_1 = 1000$	$n_2 = 2000$
sample mean	$\bar{x}_1 = 67.5$ inches	$\bar{x}_2 = 68$ inches

standard deviation = 2.5 inches

level of significance = 0.05

Step3 - Test sample statistics

$$\boxed{z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma^2\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}}$$

Step4 - Calculations

$$z = \frac{67.5 - 68}{\sqrt{(2.5)^2\left[\frac{1}{1000}+\frac{1}{2000}\right]}}$$

$$= \frac{-0.5}{0.0968}$$

$$\boxed{z = -5.16}$$

Step 5 - Conclusion.

$|z_{cal}| = 5.16$

$z_{tab} = 1.96$

$$\boxed{z_{cal} > z_{tab}}$$

∴ Hence the null hypothesis H_0 is rejected at 5% level of significance and we conclude that the samples are not drawn from the population of S.D 2.5 inches

35. A random sample of 400 students is found to have a mean height of 171.38 cms. Can it be reasonbly regarded as a sample from a large population with mean ht 171.17 cms and S.D 3.30 cms (Test at $\alpha = 5\%$)

Solution :-

Step1 - Setup for statistical hypothesis

μ = population mean.

Null hypothesis, $H_0 \Rightarrow \mu = 171.17$ cms

Alternative hypothesis, $H_1 \Rightarrow \mu \neq 171.17$ cms

Step 2 - Sample data

n = sample size = 400

$\bar{x}$ = sample mean = 171.38 cm

σ = standard deviation = 3.30

μ = population mean = 171.17 cms

α = level of significance = 0.05

Z_{tab} = 1.96

Step 3 - Test sample statistics

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

Step 4 - Calculations

$$Z = \frac{171.17 - 171.38}{\frac{3.30}{\sqrt{400}}}$$

$$= \frac{-0.21}{\frac{3.30}{20}}$$

$$\boxed{Z = -1.2727}$$

Step 5 - Conclusion

$|Z_{cal}| = 1.2727$

$Z_{tab} = 1.96$

$$\boxed{Z_{cal} < Z_{tab}}$$

∴ Hence the null hypothesis Ho is accepted at 5% level of significance and we can conclude that the sample is from a large population with mean height 171.17 cms

(36) A list of breaking strengths of two different types of cables was conducted using samples of $n_1 = n_2 = 100$ pieces of each type of cable

cable I	cable II
$\bar{x} = 1925$	$\bar{y} = 1905$
$\sigma_1 = 40$	$\sigma_2 = 30$

Do the data provide sufficient evidence to indicate a difference between the mean breaking strengths of two cables at 10% level of significance?

Solution:-

Step 1 - Setup for statistical hypothesis

μ_1 and μ_2 are two population means.

Null hypothesis, $H_0 \Rightarrow \mu_1 = \mu_2$

Alternative hypothesis, $H_1 \Rightarrow \mu_1 \neq \mu_2$

Step 2 - Sample data

	cable I	cable II
sample mean	$\bar{x} = 1925$	$\bar{y} = 1905$
S.D	$\sigma_1 = 40$	$\sigma_2 = 30$
sample size	$n_1 = 100$	$n_2 = 100$

$\alpha = 10\% = 0.10$

$Z_{tab} = 1.65$

Step 3 - Test sample statistics.

$$\boxed{Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}}$$

Step 4 - Calculations

$$Z = \frac{1925 - 1905}{\sqrt{\frac{(40)^2}{100} + \frac{(30)^2}{100}}}$$

$$= \frac{20}{5}$$

$$\boxed{Z = 4}$$

Step 5 - Conclusion

$z_{cal} = 4$

$z_{tab} = 1.65$

$\boxed{z_{cal} > z_{tab}}$

∴ Hence, the null hypothesis H_0 is rejected at $10\% = \alpha$ and we can conclude that there is no difference between the mean breaking strength of two cables.

31 Random sample of 1200 households from one town gives the mean income as Rs.500 per month with standard deviation of Rs.70 and a sample of 1000 households from another town gives the mean income as Rs600 per month, with a S.D of Rs.90. Test whether the mean income of households from two towns differ significantly or not at 0.05 level?

Solution:-

Step 1 - Setup for statistical hypothesis

Null hypothesis, $H_0 \Rightarrow \mu_1 = \mu_2$

Alternative hypothesis, $H_1 \Rightarrow \mu_1 \neq \mu_2$

Step 2 - Sample data.

	Town I	Town II
Sample size	$n_1 = 1200$	$n_2 = 1000$
sample mean	$\bar{x} = 500$	$\bar{y} = 600$
S.D	$\sigma_1 = 70$	$\sigma_2 = 90$

$\alpha = 5\% = 0.05$

$z_{tab} = 1.96$

Step 3 - Test Statistic

$$\boxed{z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}}$$

Step 4 - Calculations

$$z = \frac{500 - 600}{\sqrt{\frac{(70)^2}{1200} + \frac{(90)^2}{1000}}}$$

$$z = \frac{-100}{\sqrt{\frac{49}{12}+\frac{81}{10}}}$$

$$= \frac{-100}{\sqrt{12.1833}}$$

$$= \frac{-100}{3.4904}$$

$$\boxed{z = -28.6491}$$

Step 5- Conclusion.

$|z_{cal}| = 28.6491$

$z_{tab} = 1.96$

$\therefore \boxed{z_{cal} > z_{tab}}$

Hence, the null hypothesis H_0 is rejected and we can conclude that mean income of households do not differ significantly.

38 Test the significance of the difference between the means of the samples from the following data at 5% level:

	Size A	Sample B
size of sample	100	150
mean	50	51
S.D	4	5

Solution:-

Step 1 - Setup for statistical hypothesis,

Null hypothesis, $H_0 \Rightarrow \mu_1 = \mu_2$

Alternative hypothesis, $H_1 \Rightarrow \mu_1 \neq \mu_2$

Step 2 - Sample data

	Sample A	Sample B
sample size	$n_1 = 100$	$n_2 = 150$
mean	$\bar{x} = 50$	$\bar{y} = 51$
S.D	$\sigma_1 = 4$	$\sigma_2 = 5$

$\alpha = 5\% = 0.05$

$z_{tab} = 1.96$

Step3- Test sample statistics

$$\boxed{z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}}$$

Step4 - Calculations

$$z = \frac{50 - 51}{\sqrt{\frac{16}{100} + \frac{25}{150}}}$$

$$= \frac{-1}{\sqrt{0.16 + 0.1666}}$$

$$= \frac{1}{\sqrt{0.3266}}$$

$$\boxed{z = 1.749}$$

Step5- Conclusion

$z_{cal} = 1.749$

$z_{tab} = 1.96$

$$\boxed{z_{cal} < z_{tab}}$$

∴ Hence, the null hypothesis H_0 is accepted and we can conclude that the difference of means do differ.

(34) An insurance agent has claimed that the average age of policy holders who issue through him is less than the average for all agents which is 30.5 yrs. A random sample of 100 policy holders who had issued through him the age distribution?

Age in yrs	16-20	21-25	26-30	31-35	36-40
no. of persons	12	22	20	30	16

Test the significance 5% level. Tabulated value at 5% = 1.645

Solution :-

Age	no. of persons insured f_i	Mid value x_i	$f_i x_i$	$(x_i - x)^2$	$(x_i - x)^2 f_i$
16-20	12	18	216	116.64	1399.68
21-25	22	23	506	33.64	740.08
26-30	20	28	560	0.64	12.8
31-35	30	33	990	17.64	529.2
36-40	16	38	608	84.64	1354.24
	$\Sigma f_i = 100$		$\Sigma f_i x_i = 2880$		$\Sigma f_i (x_i - x)^2 = 4036$

$$\bar{x} = \frac{1}{N} \Sigma f_i x_i = \frac{2880}{100} = 28.8$$

$$S^2 = \frac{1}{N} \Sigma f_i (x_i - \bar{x})^2 = \frac{4036}{100} = 40.36$$

$$\boxed{S = 6.353}$$

Step 1 - Setup for statistical hypothesis

μ = averageage of policy holders

Null hypothesis, $H_0 \Rightarrow \mu = 30.5$ yrs

Alternative hypothesis, $H_1 \Rightarrow \mu < 30.5$ yrs

$\therefore$ therefore, it is a left-tailed test.

Step 2 - Sample data.

$\bar{x}$ = sample mean = 28.8

σ = S.D = 6.353

$\alpha = 5\% = 0.05$

$z_{tab} = -1.645$

Step 3 - test sample statistic

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

Step 4 - Calculations

$$z = \frac{28.8 - 30.5}{\frac{6.353}{\sqrt{100}}}$$

$$= \frac{-1.7}{0.6353}$$

$$\boxed{z = -2.676}$$

Step 5 - Conclusion

$|z_{cal}| = 2.676$

$z_{tab} = 1.645$

$$\boxed{z_{cal} > z_{tab}.}$$

∴ Hence, the null hypothesis H_0 is rejected and we can conclude that the random sample of 100 policy holders who had insured through him has an average of policy holders less than 30.5 years.

(40) A machine puts out 16 imperfect articles in a sample of 500. After machine is overhauled, it puts out 3 imperfect articles in a batch of 100. Has the machine improved?

Solution:-

Step 1 - Setup for statistical hypothesis

P_1 = proportion of imperfect articles before hauling.

P_2 = proportion of imperfect articles after hauling.

Null hypothesis, $H_0 \Rightarrow P_1 = P_2$

Alternative hypothesis, $H_1 \Rightarrow P_1 > P_2$

Step 2 - Sample data

n_1 = sample size before machine hauling = 500

n_2 = sample size articles after hauling = 100

x_1 = no. of imperfect articles before = 16

x_2 = no. of imperfect articles after hauling = 3

Step 3 - Test statistics

$$\boxed{z = \frac{p_1 - p_2}{\sqrt{pq\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}}$$

Step 4 - Calculations

$$p_1 = \frac{x_1}{n_1} = \frac{16}{500} = 0.032$$

$$p_2 = \frac{x_2}{n_2} = \frac{3}{100} = 0.03$$

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{16+3}{500+100} = \frac{19}{600} = 0.032$$

$$q = 1 - p = 0.968$$

So $$z = \frac{0.032 - 0.03}{\sqrt{0.032 \times 0.968\left[\frac{1}{500}+\frac{1}{100}\right]}}$$

$$= \frac{0.002}{0.019}$$

$$\boxed{z = 0.104}$$

Step 5 - Conclusion

$z_{cal} = 0.104$

$z_{tab} = 1.645$

$$\boxed{z_{tab} > z_{cal}}$$

∴ Thus, H0 Null hypothesis is accepted and we can conclude that the machine has improved.

Generated by CamScanner

41) In a sample of 1000 people in a state, 540 are rice eaters and the rest are wheat eaters. Can we assume that both rice and wheat are equally popular in this state at 1% level of significance? ($z_{tab} = 2.58$)

Solution:-

Step 1 - Setup for statistical hypothesis

P = Population proportion of rice eaters = 0.5

Q = Population proportion of wheat eaters = 0.5

Null hypothesis, $H_0 \Rightarrow P = 0.5$

Alternative hypothesis, $H_1 \Rightarrow P \neq 0.5$

∴ Hence, it is a two tailed test.

Step 2 - Sample data

n = population size = 1000

$\bar{x}$ = sample proportion = 540

$$p = \frac{\bar{x}}{n} = \frac{540}{1000} = 0.54$$

$$q = 1 - p = 0.46$$

α = level of significance = 0.01

$$z_{tab} = 2.58$$

Step 3 - Test sample statistics

$$\boxed{z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}}$$

Step 4 - Calculations

$$z = \frac{0.54 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{1000}}}$$

$$\boxed{z = 2.532}$$

Step 5 - Conclusion

$z_{tab} = 2.58$; $z_{cal} = 2.532$

$$\boxed{z_{cal} < z_{tab}}$$

∴ Hence, the null hypothesis is accepted and we conclude that both rice and wheat eaters are equally popular in the state at 1% level of significance.

(42) Twenty people were attacked by a disease and only 18 survived. Will you reject the hypothesis that the survival rate, if attacked by this disease, is 85% in favour of hypothesis that it is more, at 5% level.

Solution:-

Step 1- Setup for statistical hypothesis

P- Population of diseased people.

Null hypothesis, $H_0 \Rightarrow P = 0.85$

Alternative hypothesis, $H_1 \Rightarrow P > 0.85$

∴ Hence, this is a right tailed test.

Step 2- Sample data

n = sample size = 20

$\bar{x}$ = No. of survived people = 18

p = Proportion of survived people = $\frac{x}{n} = \frac{18}{20} = 0.9$

$P = 0.85$

$Q = 1 - P = 0.15$

$\alpha = 0.05$

$z_{tab} = 1.645$

Step 3- Test sample statistics

$$\boxed{z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}}$$

Step 4- Calculations

$$z = \frac{0.9 - 0.85}{\sqrt{\frac{0.85 \times 0.15}{20}}} = \frac{0.5}{0.8} = 0.625$$

$$\boxed{z = 0.625}$$

Step 5- Conclusion

$z_{tab} = 1.645$

$z_{cal} = 0.625$

$$\boxed{z_{cal} < z_{tab}}$$

∴ Hence, we accept Null hypothesis H_0 and we can conclude that the proportion of survived people is 0.85

Generated by CamScanner

[43] Before an increase in excircise duty on tea, 800 persons out of a sample of 1000 were found to be tea drinkers. After an increase in excise duty, 800 persons were tea drinkers in a sample of 1200. State whether there is a significant decrease in consupmtion of tea after the increase in execise duty? ($\alpha = 5\%$)

Solution:-

Step 1 - Setup for statistical hypothesis

Let P_1 and P_2 be the population proportions of people who drink tea before and after execise duty respectively.

Null hypothesis, $P_1 = P_2$

Alternative hypothesis, $P_1 > P_2$

Step 2 - Sample data

	Before execise duty	After execise duty
sample size	$n_1 = 1000$	$n_2 = 1200$
sample mean	$x_1 = 800$	$x_2 = 800$

Step 3 - Test Sample Statistics

$$\boxed{z = \frac{p_1 - p_2}{\sqrt{pq\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}}$$

Step 4 - Calculations

$$p_1 = \frac{x_1}{n_1} = \frac{800}{1000} = 0.8$$

$$p_2 = \frac{x_2}{n_2} = \frac{800}{1200} = 0.66$$

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{1000 \times 0.8 + 1200 \times 0.66}{1000 + 1200} = 0.7272$$

$$q = 1 - p = 0.2727$$

$$z = \frac{0.8 - 0.66}{\sqrt{\frac{0.7272 \times 0.2727}{1}\left[\frac{1}{1000} + \frac{1}{1200}\right]}} = \frac{0.14}{\sqrt{3.6356 \times 10^{-4}}}$$

$$\boxed{z = 7.3424}$$

Step 5 Conclusion

Since $\boxed{Z_{cal} > Z_{tab}}$, we reject the Null hypothesis H_0 and conclude that there is a significant difference in the consumption of tea before and after an increase on excise duty on tea.

(44) A random sample of 500 apples was taken from a large consignment and 60 were found bad. Obtain the 98% confidence limits for the percentage of bad apples in consignment ($z = 2.33$)?

Solution:-

Step 1 Setup for statistical hypothesis

Null hypothesis, $H_0 \Rightarrow P = P_0$

Alternative hypothesis, $H_1 \Rightarrow P \neq P_0$

Step 2 - Sample data.

n = sample size = 500

p = proportion of bad apples $= \frac{60}{500} = 0.12$

$q = 1 - p = 0.88$

Step 3 - Limits of population proportion

$$\Rightarrow \boxed{p \pm z\sqrt{\frac{pq}{n}}}$$

Step 4 - Calculations

$$p \pm 3\sqrt{\frac{pq}{n}} = 0.12 \pm (2.33)\sqrt{\frac{0.12 \times 0.88}{500}}$$

$$= 0.12 \pm (2.33)\sqrt{2.112 \times 10^{-4}}$$

$$= 0.12 \pm 0.03386$$

$$\boxed{= (0.08614, 0.15386)}$$

Step 5 - Conclusion

The percentage of bad apples in the consignment lies between 8.6 and 15.38

15. A coin is tossed 10,000 times and it turns up head 5195. Test the hypothesis that the coin is unbiased at 1% level.

Solution:-

Step 1- Setup for statistical hypothesis

Null hypothesis, $H_0 \Rightarrow P = 0.5$

Alternative hypothesis, $H_1 \Rightarrow P \neq 0.5$

$\therefore$ Hence, it is a two tailed test.

Step 2- Sample data.

$n = 10,000$ = sample size

$x = 5195$ - No. of times head turns up

$$p = \frac{x}{n} = \frac{5195}{10000} = 0.5195$$

$q = 1 - p = 1 - 0.5195 = 0.4805$

$\alpha = 1\% = 0.01$

$z_{tab} = 2.58$

Step 3- Test sample statistics

$$\boxed{z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}}$$

Step 4- Calculations

$$z = \frac{0.5195 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{10000}}}$$

$$\boxed{z = 3.9}$$

Step 5- Conclusion

$z_{tab} = 2.58$

$z_{cal} = 3.9$

$$\boxed{z_{cal} > z_{tab}}$$

$\therefore$ Hence, the null hypothesis H_0 is rejected and we can conclude that the coin is biased.

46. A person throws a dice 9000 times 3240 times 3 or 4. Can you conclude the dice is unbiased one at 5%?

Solution:-

Step 1 - Setup for statistical hypothesis

Null hypothesis, $H_0 \Rightarrow P = 0.333$

Alternative hypothesis $\Rightarrow P \neq 0.333$

Step 2 - Sample data

n = sample size = 9000

x = no. of times 3 or 4 obtained = 3240

$$p = \frac{x}{n} = \frac{3240}{9000} = 0.36$$

$$q = 1 - p = 0.64$$

$$P = P(\text{getting } 3 \text{ or } 4) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3} = 0.333$$

$$Q = 1 - P = 0.6667$$

Step 3 - Sample Test statistics

$$\boxed{z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}}$$

Step 4 - Calculations

$$z = \frac{0.36 - 0.333}{\sqrt{\frac{0.333 \times 0.6667}{9000}}}$$

$$= \frac{0.027}{4.96 \times 10^{-3}}$$

$$\boxed{z = 5.38}$$

Step 5 - Conclusion

$z_{tab} = 1.96$ $\boxed{z_{cal} > z_{tab}}$

$z_{cal} = 5.38$

∴ Hence, the null hypothesis is rejected and we can conclude that the die is biased.

Generated by CamScanner

(41) 1000 articles from a factory A are examined and found to have 3% defectives. 1500 similar articles from factory B are 2% defectives. Can it be resonably concluded that the product of the first factory is inferior to the second?

Solution:-

Step 1 - Set up for statistical hypothesis.

Null hypothesis, $H_0 \Rightarrow P_1 = P_2$

Alternative hypothesis, $H_1 \Rightarrow P_1 \neq P_2,\ P_1 > P_2$

$\therefore$ Since $P_1 > P_2$, it is a right tailed test.

Step 2 - Sample data

	Factory A	Factory B
Sample size	$n_1 = 1000$	$n_2 = 1500$
sample attributes	$P_1 = 0.03$	$P_2 = 0.02$

Step 3 - Test sample statistics

$$\boxed{z = \frac{P_1 - P_2}{\sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}}}$$

Step 4 - Calculations

$Q_1 = 1 - P_1 = 1 - 0.03 = 0.97$

$Q_2 = 1 - P_2 = 1 - 0.02 = 0.98$

$$z = \frac{0.03 - 0.02}{\sqrt{\frac{0.03 \times 0.97}{1000} + \frac{0.02 \times 0.98}{1500}}}$$

$$z = \frac{0.01}{\sqrt{\frac{0.0291}{1000} + \frac{0.0196}{1500}}} = \frac{0.01}{6.4935 \times 10^{-3}} = \boxed{1.54}$$

Step 5 - Conclusion.

$z_{tab} = 1.65$ (at 5% of level of significance)

$z_{cal} = 1.54$

$\boxed{z_{cal} < z_{tab}}$

$\therefore$ Hence, H_0 is accepted and hence, we conclude that the product of first factory is inferior to the second.

(48) A sample of 20 items has mean 42 units and S.D 5 units. Test the hypothesis that it is a random sample from a normal population with mean 45 units.

Solution :-

Step 1 - Setup for statistical hypothesis,

Null hypothesis, $H_0 \Rightarrow \mu = 45$

Alternative hypothesis, $H_1 \Rightarrow \mu \neq 45$

Step 2 - Sample data

n = sample size = 20

$\bar{x}$ = sample mean = 42

μ = population mean = 45

$\sigma = 5$

$\alpha = 5\%$

$t_{tab} = 2.093$

Step 3 - Test sample statistics

$$\boxed{t = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n-1}}}}$$

Step 4 - Calculations

$$t = \frac{42 - 45}{\frac{5}{\sqrt{20-1}}}$$

$$\boxed{t = -2.615}$$

Step 5 - Conclusion

$t_{tab} = 2.093$

$|t_{cal}| = 2.615$

$$\boxed{t_{cal} > t_{tab}}$$

$\therefore$ Hence, the null hypothesis H_0 is rejected and we can conclude that the sample could not have come from this population.

(49) A sample of 10 boys has the I.Qs 70, 120, 110, 101, 88, 83, 95, 98, 107 and 100. Test the mean IQ of students is 100 at 0.05 level of significance.

Solution :-

Step 1 - Setup for statistical hypothesis

Null hypothesis, $H_0 \Rightarrow \mu = 100$

Alternative hypothesis, $H_1 \Rightarrow \mu \neq 100$

Step 2 - Sample data.

Since, the S.D and mean of sample are not given directly.

x	$x-\bar{x}$	$(x-\bar{x})^2$
70	-27.2	739.84
120	22.8	519.84
110	12.8	163.84
101	3.8	14.44
88	-9.2	84.64
83	-14.2	201.64
95	-2.2	4.84
98	0.8	0.64
107	9.8	96.04
100	2.8	7.84
$\Sigma x = 972$		$\Sigma(x-\bar{x})^2 = 1833.60$

Mean, $\bar{x} = \frac{\Sigma x}{n} = \frac{972}{10} = 97.2$

we know that, $S^2 = \frac{1}{n-1} \Sigma(x_i - \bar{x})^2 = \frac{1883.60}{9}$

$\therefore$ Standard Deviation, $S = \sqrt{203.73} = 14.27$

Step 3 - Test sample statistics

$$\boxed{t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}}$$

Step 4 - Calculations

$$t = \frac{97.2 - 100}{14.27/\sqrt{10}}$$

$$\boxed{t = -0.62}$$

Step 5 - Conclusion.

$|t_{cal}| = 0.62$

$t_{tab} = 2.262$

$$\boxed{t_{cal} < t_{tab}}$$

∴ Hence H_0 is accepted and we can conclude that the data supports the population mean IQ is 100.

50. The heights of 10 males in a locality are found to be 70, 67, 62, 68, 61, 68, 70, 64, 64, 66 inches. Is it resonable to belive that the average height is greater than 64 inches? Test at 5% significance level assuming that for 9 degrees of freedom ($t = 1.833$ at $\alpha = 0.05$)

Solution:

Step 1 - Setup for statistical hypothesis

Null hypothesis, $H_0 \Rightarrow$ $\mu = 64$ inches

Alternative hypothesis, $H_1 \Rightarrow \mu_1 > 64$ inches

Step 2 - Sample data

Sample size = 10

$\alpha = 0.05$

degree of freedom = 9

Since, the S.D and mean of sample are not given directly, they are calculated.

x	$x-\bar{x}$	$(x-\bar{x})^2$
70	4	16
67	1	1
62	-4	16
68	2	4
61	-5	25
68	2	4
70	4	16
64	-2	4
64	-2	4
66	0	0
$\Sigma x = 660$		$\Sigma(x-\bar{x})^2 = 90$

we know that $S^2 = \frac{1}{n-1} \Sigma (x_i - \bar{x})^2 = \frac{1}{9} \times 90 = 10$

$\therefore$ sample S.D $= \sqrt{10} = 3.16$

Step 3 - Test statistics

$$\boxed{t = \frac{\bar{x} - \mu}{S/\sqrt{n-1}}}$$

Step 4 - Calculations

$$t = \frac{66-64}{3.16/\sqrt{9}} = \frac{2}{1.05} = 1.9$$

$\boxed{t = 1.9}$

Step 5 - Conclusion

$t_{cal} = 1.9$

$t_{tab} = 1.833$

$\boxed{t_{cal} > t_{tab}}$

$\therefore$ Hence, H_0 is rejected and we can conclude that the average height is greater than 64.

51. Two independent samples of 8 and 7 respectively had the following values.

Sample I	11	11	13	11	15	9	12	14
Sample II	9	11	10	13	9	8	10	-

Is there difference between the means of samples significant? Test at 1% level of significance.

Solution :

Step 1 - Setup for statistical hypothesis

Null hypothesis, $H_0 \Rightarrow \mu_1 = \mu_2$

Alternative hypothesis, $H_1 \Rightarrow \mu_1 \neq \mu_2$ (two-tailed test)

Step 2 - Sample data.

	A	B
sample size	8	7
sample mean	$\bar{x} = \frac{\Sigma x}{n_1} = \frac{96}{8} = 12$	$\bar{y} = \frac{\Sigma y}{n_2} = \frac{70}{7} = 10$

x	$x-\bar{x}$	$(x-\bar{x})^2$	y	$(y-\bar{y})$	$(y-\bar{y})^2$
11	-1	1	9	-1	1
11	-1	1	11	1	1
13	1	1	10	0	0
11	-1	1	13	3	9
15	3	9	9	-1	1
9	-3	9	8	-2	4
12	0	0	10	0	0
14	2	4			
$\Sigma x = 96$		$\Sigma(x-\bar{x})^2 = 26$	$\Sigma y = 70$		$\Sigma(y-\bar{y})^2 = 16$

Now, $s^2 = \frac{1}{n_1+n_2-2} \left[\Sigma(x_i-\bar{x})^2 + \Sigma(y_i-\bar{y})^2 \right]$

$= \frac{1}{8+7-2} (26+16) = \frac{42}{13} = 3.23$

$S = 1.8$

Step 3 - Test statistics

$$t = \frac{\bar{x} - \bar{y}}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Step 4 - Calculations

$$t = \frac{12 - 10}{1.8\sqrt{\frac{1}{8} + \frac{1}{7}}}$$

$$t = \frac{2}{0.9316} = \boxed{2.15}$$

Step 5 - Conclusion

$t_{tab} = 3.012$

$t_{cal} = 2.15$

$\boxed{t_{tab} > t_{cal}}$

∴ Hence, Ho is accepted and we conclude that the difference between the means is significant.