Probability and statistics

3130006 [Paper solution]

Date 10 10 20

Q:1

(a) In how many different ways can 4 of 15 laboratory assistant be chosen to assist with an experiment?

→ Total assistant = 15

laboratory assistant to be selected = 4

- no. of ways selecting 4 laboratory assistants with an experiment

$$= {}^{15}C_4$$

$$= \frac{15 \times 14 \times 13 \times 12}{4 \times 3 \times 2 \times 1}$$

$$= \boxed{1365}$$

(b) If 5 of 20 tires in storage are defective and 5 of them are randomly chosen for inspection (i.e each tire has same chance of being selected) what is the prob. that the two of the defective tire will be selected?

→ Let X denote the number of defective tire in a draw of 5 from group of 20

Here Total tire are 20 = n

Prob. of defective tire $= \frac{5}{20} = \frac{1}{4}$

Prob. of non-defective tire $= \frac{15}{20} = \frac{3}{4}$

- Prob. that two of defective tire will be selected

= Atleast two be will be selected

$$P(X \geq 2) = P(X=2) + P(X=3) + P(X=4) + P(X=5)$$
$$= 1 - [P(X=0) + P(X=1)]$$

- $P(X=0) = \dfrac{\binom{5}{0}\binom{15}{5}}{\binom{20}{5}} = \dfrac{15 \times 14 \times 13 \times 12 \times 11}{20 \times 19 \times 18 \times 17 \times 16} = 0.1937$

- $P(X=1) = \dfrac{\binom{5}{1}\binom{15}{4}}{\binom{20}{5}} = \dfrac{15 \times 14 \times 13 \times 12 \times 5 \times 5 \times 4 \times 3 \times 2 \times 1}{20 \times 19 \times 18 \times 17 \times 16 \times 4 \times 3 \times 2 \times 1}$
$= 0.44$

— $P(X \geq 2) = 1 - (0.1937 + 0.44)$
$\boxed{P(X \geq 2) = 0.3663}$

(c) The following are the data on the drying time of a certain varnish and the amount of an additive that is intended to reduce the drying time :

Amount of varish (x)	0	1	2	3	4	5	6	7	8
Drying time (y)	12	10.5	10	8	7	8	7.5	8.5	9

i) fit second degree polynomial by least square
ii) use the result i) to predict the drying time of varnish when 6.5 gm of the additive is being used.

→ second degree polynomial is $y = a + bx + cx^2$
$n = 9$

	x_i	y_i	x_i^2	x_i^3	x_i^4	x_iy_i	$x_i^2y_i$
	0	12	0	0	0	0	0
	1	10.5	1	1	1	10.5	10.5
	2	10	4	8	16	20	40
	3	8	9	27	81	24	72
	4	7	16	64	256	28	112
	5	8	25	125	625	40	200
	6	7.5	36	216	1296	45	270
	7	8.5	49	343	2401	59.5	416.5
	8	9	64	512	4096	72	576
Σ	36	80.5	204	1296	8772	299	1697

- $$\sum_{i=0}^{8} y_i = an + b\sum_{i=0}^{8} x_i + c\sum_{i=0}^{8} x_i^2$$

$$80.5 = 9a + 36b + 204c \quad \text{—— (1)}$$

- $$\sum_{i=0}^{8} x_iy_i = a\sum_{i=0}^{8} x_i + b\sum_{i=0}^{8} x_i^2 + c\sum_{i=0}^{8} x_i^3$$

$$299 = 36a + 204b + 1296c \quad \text{—— (2)}$$

- $$\sum_{i=0}^{8} x_i^2y_i = a\sum_{i=0}^{8} x_i^2 + b\sum_{i=0}^{8} x_i^3 + c\sum_{i=0}^{8} x_i^4$$

$$1697 = 204a + 1296b + 8772c \quad \text{—— (3)}$$

$\longrightarrow$ By solving 1, 2 & 3

$a = 12.1847$

$b = -1.8465$

$c = 0.1829$

- $y = a + bx + cx^2$

$$\boxed{y = 12.1847 - 1.8465x + 0.1829x^2}$$

ii) for additive used for varnish is 6.5 then drying time is

i.e for $x = 6.5$ $y = (?)$

$y = 12.1847 - 1.8465x + 0.1829x^2$

$y = 12.1847 - 1.8465(6.5) + 0.1829(6.5)^2$

$y = 7.909975$

$\boxed{y \approx 7.91}$

Q:2

ⓐ If 3 balls are randomly drawn from a bowl containing 6 white and 5 Black balls what is the probability that one of the ball is white and other two are black?

→ Here 3 ball are selected

Total balls are $6 + 5 = 11$ balls

white ball $= 6$

Black ball $= 5$

- Prob. that one ball selected is white and other two are black

= (Selection of white ball) (selection of Black ball) / (Total balls)

$$= \frac{\binom{6}{1}\binom{5}{2}}{\binom{11}{3}} = \frac{6 \times 5 \times 4 \times 3 \times 2}{2 \times 11 \times 10 \times 9}$$

$= \boxed{4/11}$

Page No.
Date

(5) The article "A Thin film oxygen uptake test for the evaluation of automotive crankcase lubricants" reported the following data on ~~oxygen~~ oxygen oxidation - induction time (min) for various commercial oils:
87, 103, 130, 160, 180, 195, 132, 145, 211, 105, 145, 153, 152, 138, 87, 99, 93, 119, 129

i) Calculate the sample variance and standard deviation.
ii) If the observation were re-expressed in hours what would be the resulting values of sample variance and standard deviation.

i)

x_i	$(x_i-\bar{x})^2$
87	2293.93
103	1017.29
130	23.96
160	630.26
180	2034.46
195	3612.61
132	8.38
145	102.11
211	5791.97
105	893.71
145	102.11
153	327.79
152	292.58
138	9.64
87	2293.93
99	1288.45
93	1755.19
119	252.65
129	34.75
Σ 2563	22765.77

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{2563}{19}$$

$$\bar{x} = 134.895$$

$$S.d. = \sqrt{\frac{\Sigma(x_i-\bar{x})^2}{n}} = \sqrt{\frac{22765.77}{19}}$$

$$\boxed{S.D. = 34.615 \text{ min}}$$

$$\text{variance} = \frac{\Sigma(x_i-\bar{x})^2}{n} = \frac{22765.77}{19}$$

$$\boxed{\text{variance} = 1198.198 \text{ min}^2}$$

ii) when it is reexpressed in hours
1 hr = 60 min

$\therefore$ sample new variance will be $\frac{1198.198}{3600}$

variance $= \boxed{0.3328\ hr^2}$

$\therefore$ sample new S.D. will be $\frac{34.615}{60}$

$\boxed{S.D. = 0.5769\ hr}$

(c) In an examination minimum 40 marks for passing and 75 marks for distinction are required. In this examination 45% students passed and 9% obtained distinction. find average marks and S.D. of this distribution of marks
$P(Z = 0.125) = 0.05$ and $P(z = 1.34) = 0.41$

$\rightarrow$ Let 'μ' be the average marks (mean) and 'σ' be the standard Deviation of normal curve

$P(X < 40) = 0.45$

$P(X > 75) = 0.09$

since $P(X < 40) < 0.5$ so the corresponding value of 'z' will be -ve,

$X = 40$; $z = \frac{40 - \mu}{\sigma} = -z_1$

$P(X > 75) < 0.5$ so corresponding value of z will be "+ve"

$x = 75$; $z = \frac{75-\mu}{\sigma} = z_2$

- $P(z < -z_1) = 0.45$

$0.5 - P(0 < z < z_1) = 0.45$

$P(0 < z < z_1) = 0.05$

$z_1 = P(0.05)$ $\{$ Given $P(0.05) = 0.125 \}$

$z_1 = 0.125$

$\therefore \frac{40-\mu}{\sigma} = -0.125$ — (1)

- $P(z > z_2) = 0.09$

$0.5 - P(0 < z < z_2) = 0.09$

$P(0 < z < z_2) = 0.41$

$z_2 = P(0.41)$ $\{$ Given $P(0.41) = 1.34 \}$

$z_2 = 1.34$

$\therefore \frac{75-\mu}{\sigma} = 1.34$ — (2)

- By solving 1 & 2

$40 - \mu = -0.125\sigma$

$75 - \mu = 1.34\sigma$

$- \quad + \quad -$

$35 = 1.465\sigma$ $\boxed{\sigma = 23.89}$

$\boxed{\mu = 42.9874}$

$\therefore$ The average marks is 42.9874 and S.D is 23.89

Page No.
Date

(c) Distribution of height of 1000 students is normal with mean 165 cms and S.D. 15 cms How many soldiers are of height.

i) less than 138 cm

ii) more than 198 cm

iii) between 138 and 198 cm

[$P(Z=1.8) = 0.4641$, $P(Z=2.2) = 0.4861$]

→ Let random variable 'X' represent the heights of students that follow normal distribution with mean 165cm and S.D. of 15 cms

i) less than 138 cm

$$P(X \leq 138) = P\left(\frac{X-\mu}{\sigma} \leq \frac{138-165}{15}\right)$$

$$= P(Z \leq -1.8)$$

$$= 0.5 - P(0 < Z < 1.8)$$

$$= 0.5 - 0.4641$$

$$\boxed{P(X \leq 138) = 0.0359}$$

ii) more than 198 cm

$$P(X \geq 198) = P\left(\frac{X-\mu}{\sigma} \geq \frac{198-165}{15}\right)$$

$$= P(Z \geq 2.2)$$

$$= 0.5 - P(0 < Z < 2.2)$$

$$= 0.5 - 0.4861$$

$$\boxed{P(X \geq 198) = 0.0139}$$

iii) between 138 and 198 cm

$Z_1 = \frac{X-\mu}{\sigma} = \frac{138-165}{15}$, $Z_2 = \frac{X-\mu}{\sigma} = \frac{198-165}{15}$

$Z_1 < 0$, $Z_2 > 0$ $Z_1 = -1.8$, $Z_2 = 2.2$

$P(-Z_1 \leq X \leq Z_2)$

$= P(-Z_1 \leq Z \leq 0) + P(0 \leq Z \leq Z_2)$

$= P(0 \leq Z \leq Z_1) + P(0 \leq Z \leq Z_2)$

$= P(0 \leq Z \leq 1.8) + P(0 \leq Z \leq 2.2)$

$= 0.4641 + 0.4861$

$= \boxed{0.9502}$

Q:3

(a) Compute the coefficient of correlation between X and Y using the following data :

X	2	4	5	6	8	11
Y	18	12	10	8	7	5

→ coefficient of correlation $\gamma = \frac{\Sigma X'Y'}{\sqrt{\Sigma (X')^2 \Sigma (Y')^2}}$

here $X' = X - \bar{X}$; $Y' = Y - \bar{Y}$

	X	Y	$X' = X - \bar{X}$	$Y' = Y - \bar{Y}$	$(X')^2$	$(Y')^2$	$X'Y'$
	2	18	-4	8	16	64	-32
	4	12	-2	2	4	4	-4
	5	10	-1	0	1	0	0
	6	8	0	-2	0	4	0
	8	7	2	-3	4	9	-6
	11	5	5	-5	25	25	-25
Σ	36	60	0		50	106	-67

$\bar{x} = \frac{\Sigma x_i}{n} = \frac{36}{6}$ $\boxed{\bar{x} = 6}$

$\bar{y} = \frac{\Sigma y_i}{n} = \frac{60}{6}$ $\boxed{\bar{y} = 10}$

$\therefore \quad r = \frac{-67}{\sqrt{(50)(106)}}$

$\boxed{r = -0.92}$

(b) An analysis of monthly wages paid to workers in two firms A and B belongs to the same industry gave the following results.

	Firm A	Firm B
No. of wages earners	986	548
Average monthly wages	Rs 52.5	Rs 47.5
variance	100	121

a) which firms pays out large amount as wage bill ?

b) In which firm there is greater variability in individual wages ?

→

a) no. of wage earner in firm A = 986

Mean of monthly wage of firm A = 52.5 Rs

mean of monthly age of firm A = $\frac{\text{Total amount Paid}}{\text{no. of wage earner in firm A}}$

$52.5 \cancel{986} = \frac{\text{Total amount Paid}}{986}$

$\therefore$ Total amount paid by firm A = 51765

ii) no. of wage earner in firm B = 548
mean of monthly wage of firm B = 47.5 Rs

mean of monthly wage of firm B = $\frac{\text{Total amount Paid}}{\text{no. of wage}}$

$$47.5 = \frac{\text{Total amount paid}}{548}$$

$\therefore$ Total amount Paid by firm B = 26030

so finally from above we can say that firm A pays out large amount as wage bill which is 51765

b)

Here the variance of firm A $(\sigma^2) = 100$
$\therefore$ standard deviation $(\sigma) = 10$

variance of firm B $(\sigma^2) = 121$
$\therefore$ standard deviation $(\sigma) = 11$

Here whose standard deviation is more will have greater variability in individual wages

$\therefore$ Here the firm B will have greater variability in individual wages.

Page No.
Date

Q:5

(a) A sample of 20 items has mean 42 units and standard deviation 5 units. Test the hypothesis that its random sample from normal population with mean 45 units [at 5% level of significance for 19 d.f is 2.09]

→ $n = 20$, $\bar{x} = 42$, $S.D = 5$, $\mu = 45$

Applying t-test

i) Null Hypothesis H_0 : $\mu = 45$ (i.e mean is equal)
ii) Alternative Hypothesis H_1 : $\mu \neq 45$ (Two tailed test)
iii) Level of significance $\alpha = 0.05$
iv) Test statistics :

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

here $s^2 = \frac{n}{n-1} S^2 = \frac{20}{20-1} (5)^2 =$

$\therefore s^2 = 26.31$

$\therefore s = 5.129$

$$t = \frac{42-45}{5.129/\sqrt{20}} = -2.6157$$

$|t| = 2.6157$

v) critical value :

$\nu = n - 1 = 20 - 1 = 19$

$t_{0.05} \ (\nu = 19) = 2.09$

vi) Decision :

$|t| > t_{0.05}$

$\therefore$ Null Hypothesis is rejected

i.e there is difference between sample mean and population mean

(b) A university warehouse has received a shipment of 25 printers of which 10 are laser printers and 15 are inkjet model. If 6 of these 25 are selected randomly to check by particular technician what is prob. that exactly three of these are laser printers (so that other three are inkjet) ?

→ Total printers = 25

Laser printers = 10

Inkjet printers = 15

~~Prob. of laser printer~~

6 printers are selected

Prob. of exactly three laser printer to be selected which means three laser and three inkjet printer

$$P(X=3) = \frac{\binom{10}{3}\binom{15}{3}}{\binom{25}{6}}$$

$$= \frac{10 \times 9 \times 8 \times 15 \times 14 \times 13 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 3 \times 2 \times 25 \times 24 \times 23 \times 22 \times 21 \times 20}$$

$\boxed{P(X=3) = 0.3083}$ = Prob. of exactly three laser printer to be selected q:4

(a) Each sample of water has 10% chance of containing a particular organic pollutant. Assume that the sample are independent with regard to the presence of the pollutant. find the prob. that in next 18 samples atleast 4 sample contain the Pollutant.

→ Let 'X' be the ~~sa~~ number of sample that contain the pollutant

Here $n = 18$ $\qquad p = \frac{10}{100} = 0.1$

$q = 1 - p = 1 - 0.1 = 0.9$

- Prob. that atleast 4 sample contains the pollutant

$P(X \geq 4) = 1 - P(X < 4)$

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

$= 1 - \left[\binom{18}{0}(0.1)^0(0.9)^{18} + \binom{18}{1}(0.1)^1(0.9)^{17} + \binom{18}{2}(0.1)^2(0.9)^{16} + \binom{18}{3}(0.1)^3(0.9)^{15} \right]$

$= 1 - [0.150 + 0.300 + 0.284 + 0.168]$

$= \boxed{0.098}$

(b) Goal scored by two teams A and B in football season were as follow:

no. of goal scored in match	0	1	2	3	4
no. of matches played by team A	27	9	8	5	4
no. of maches played by team B	17	9	6	5	3

find out which team is more consistent

→ Here the team whose Standard Deviation will be less will be more consistent

for team A

	x_i	f_i	$f_i x_i$	$(x_i-\bar{x})$	$(x_i-\bar{x})^2$	$f_i(x_i-\bar{x})^2$
	0	27	0	-1.0566	1.1164	30.1428
	1	9	9	-0.0032	0.00001024	0.00009216
	2	8	16	0.9434	0.89	7.12
	3	5	15	1.9434	3.7768	18.884
	4	4	16	2.9434	8.6636	34.6544
Σ		53	56			90.80

$$\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{56}{53} = 1.0566$$

$$\text{S.d. } (\sigma) = \sqrt{\frac{\Sigma f_i(x_i-\bar{x})^2}{n}} = \sqrt{\frac{90.80}{53}} = 1.3088$$

S.d. = 1.3088

$$\text{C.V. for team A} = \frac{\sigma}{\bar{x}} \times 100 = \frac{1.31}{1.0566} \times 100 =$$

for team B

x_i	f_i	$f_i x_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
0	17	0	-1.2	1.44	24.48
1	9	9	-0.2	0.04	0.36
2	6	12	0.8	0.64	3.84
3	5	15	1.8	3.24	16.2
4	3	12	2.8	7.84	23.52
Σ	40	48			68.4

$$\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{48}{40} = 1.2$$

$$\text{S.d.} = \sqrt{\frac{\Sigma f_i (x_i - \bar{x})^2}{n}} = \sqrt{\frac{68.4}{40}} = 1.30766$$

- so Here as the standard deviation of team B is less than team A

so team B is more consistent

(c) Out of 800 families with 4 children each, how many families would be expected to have

i) 2 girl and 2 boy

ii) atleast one boy

iii) no girl

iv) at most two girl

Assume equal probabilities for boys and girls.

→ Here ~~n = 4~~ ~~N = 800 families~~

$n = 4$ N = 800 families

Let 'p' be the prob. of having boy or girl in each family

$p = \frac{1}{2}$ $\quad q = 1-p = 1-\frac{1}{2} = \frac{1}{2}$ $\quad n = 4$

$N = 800$

- Prob. of having x boys and x girls will be

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

i) Two boys and Two girls

$$P(X=2) = \binom{4}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2$$

$$= \frac{4\times3}{2^4\times2} = \frac{3}{8}$$

Expected number of families having 2 boys and 2 girls

$$= NP(X=2)$$

$$= 800 \times \frac{3}{8}$$

$$= \boxed{300}$$

ii) Atleast one boy

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

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

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

$$= 1 - \frac{1}{2^4} = \frac{15}{16}$$

expected number of families having atleast one boy

$= N\,P(X \geq 1)$

$= 800 \times \frac{15}{16}$

$= \boxed{750}$

iii) no girl i.e all are boys $P(X=4)$

$P(X=4) = \binom{4}{4} (1/2)^4 (1/2)^0$

$= \frac{1}{2^4} = \frac{1}{16}$

Expected number of families having no girl

$= N\,P(X=4)$

$= 800 \times \frac{1}{16}$

$= \boxed{50}$

iv) Atmost two girls

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

$= \binom{4}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^4 + \binom{4}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^3 + \binom{4}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2$

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

$= \frac{11}{16}$

Expected number of families having atmost two girls

$= N\,P(X \leq 2)$

$= 800 \times \frac{11}{16}$

$= \boxed{550}$

Page No.

Date

Q:4 Assume that the prob. that a wafer contains a
(a) large particle of contamination is 0.01 and that the wafer are independent that is prob. that a wafer contain a large particle is not dependent on the characteristics of any of the other wafer If ~~$~~ 15 wafer are analyzed what is prob. that no large Particles are found?

→ let 'E' be the event that wafer contains no large particles

Here E_i ~~is~~ contains no large particles
$i = 1, 2, 3, 4, \ldots, 15$

Prob. that wafer contain large particle $= 0.01$

$P(E)$ = event that contain no particle $= 0.99$

$P(E_i) = 0.99)$ $\quad$ ~~$P(E_1 \neq E_2, E_3, \ldots E_{15}) \neq 0.99$~~

$P(E_1) = P(E_2) = P(E_3) = \ldots = P(E_{15}) = 0.99$

∴ Required prob.
$P(E_1 \cap E_2 \cap E_3 \ldots \cap E_{15}) = P(E_1)P(E_2)\ldots P(E_{15})$
{As event is independent}

$= (0.99)^{15}$
$= \boxed{0.86}$

(b)

(b) A mickochip company has two machines that produces the chips Machine I produces 65% of chips, but 5% of the chips are defective. Machine II produces 35% of chips and 15% of its chips are defective. A chip is selected at random and found to be defective what is prob. that it came from machine I?

→ Let A_1 and A_2 be the event that factory produces ~~all~~ chips by machine I and machine II respectively.

$P(A_1)$ = Prob. of machine I produces chips

$= \frac{65}{100} = 0.65$

$P(A_2)$ = Prob. of machine II produces chips

$= \frac{35}{100} = 0.35$.

- 'B' be the event that machine produces defective chips

$P(B|A_1)$ = Prob. of machine I produces defective chips

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

$P(B|A_2)$ = Prob. of machine II produce defective chips

$= \frac{15}{100} = 0.15$.

i) $P(A_1|B)$ = Prob. of given that defective ~~as~~ chip is produced by machine I

$$P(A_1|B) = \frac{P(A_1)\,P(B|A_1)}{P(A_1)\,P(B|A_1) + P(A_2)\,P(B|A_2)}$$

$$= \frac{(0.65)(0.05)}{(0.65)(0.05) + (0.35)(0.15)}$$

$$\boxed{P(A_1|B) = 0.3824}$$

(c) If a publisher of nontechnical books takes great pains to ensure that its books are free of typo- -graphical errors so that prob. of any given page containing at least one such error is 0.005 and errors are independent from page to page what is prob. that one of its 400 page novels will contain i) exactly one page with errors ii) atmost three pages with errors ?

→ let 'p' be the prob. that the given page has atleast one such error

$P(X) = p = 0.005$

$n = 400$ = Total pages of novels

- The distribution has parameter $\lambda = np$

$= 400 \times 0.005$

$= 2$

Acc. to prob. mass function of poisson distribution is

$$P(X = x) = \frac{e^{-\lambda}\,\lambda^{x}}{x!} \quad ; \quad x = 0, 1, 2, \ldots$$

$$= \frac{e^{-2}\,(2)^{x}}{x!} \quad ; \quad x = 0, 1, 2, \ldots$$

i) ~~Ea~~ Exactly one page with error

$$P(X=1) = \frac{e^{-2}(2)^1}{1!} = 0.2707$$

$$\boxed{P(X=1) = 0.2707}$$

ii) Atmost three pages with error

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

$$= \frac{e^{-2}(2)^0}{0!} + \frac{e^{-2}(2)^1}{1!} + \frac{e^{-2}(2)^2}{2!} + \frac{e^{-2}(2)^3}{3!}$$

$$= 0.1354 + 0.2707 + 0.2707 + 0.18048$$

$$= 0.85728$$

$$\boxed{P(X \leq 3) \approx 0.8573}$$

Q:5

(a) Samples of sizes 10 and 14 were taken from two normal population with standard deviation 3.5 and 5.2. The sample means were found to be 20.3 and 18.6. Test whether the means of two population are same at 5% level.
[$t_{0.05} = 2.0739$]

→ ~~A~~ Here $n_1 = 10$ $S_1 = 3.5$ $\bar{X_1} = 20.3$

$n_2 = 14$ $S_2 = 5.2$ $\bar{X_2} = 18.6$

i) Null Hypothesis H_0 : $\mu_1 = \mu_2$ i.e. the mean of two population are the same

ii) Alternative Hypothesis H_1 : $\mu_1 \neq \mu_2$ (Two tailed test)

iii) Level of significance $\alpha = 0.05$

iv) Test statistics :

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

here $S^2 = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{10(3.5)^2 + 14(5.2)^2}{10+14-2}}$

$S = 4.772$

$$t = \frac{20.3 - 18.6}{4.772\sqrt{\frac{1}{10} + \frac{1}{14}}} = \cancel{2.0078}\ 0.8604$$

$|t| = 0.8604$

v) critical value :

$\nu = n_1 + n_2 - 2$

$= 10 + 14 - 2$

$= 22$

$t_{0.05}(\nu = 22) = 2.0739$

vi) Decision :

since $|t| < t_{0.05}$

∴ i.e the null Hypothesis is accepted

∴ There is no significant difference between their means

(b) Two independent samples 8 and 7 items respectively had the following values of variable (weight in kg)

0.17143

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

Do the two estimates of population variance differ significantly ? Given that for (7, 6) d.f the value of F at 5% level of significance is 4.20 nearly

iv) Test statis

→ Here $n_1 = 8$, $n_2 = 7$

	X_{1i}	$(X_1 - \bar{X}_1)^2$	X_{2i}	$(X_2 - \bar{X}_2)^2$
	9	7.5625	10	0.1849
	11	0.5625	12	~~2.6~~ 2.4649
	13	1.5625	10	~~2.4649~~ 0.1849
	11	0.5625	14	12.7449
	15	10.5625	9	2.0449
	9	7.5625	8	5.9049
	12	0.0625	10	0.1849
	14	5.0625		
Σ	94	33.5	73	23.7143

- $\bar{X}_1 = \frac{\Sigma X_{1i}}{n} = \frac{94}{8} = 11.75$

- $\bar{X}_2 = \frac{\Sigma X_{2i}}{n} = \frac{73}{7} = 10.43$

- $S_1 = \sqrt{\frac{\Sigma (X_1 - \bar{X}_1)^2}{n_1}} = \sqrt{\frac{33.5}{8}} = 2.046$

- $S_2 = \sqrt{\frac{\Sigma (X_2 - \bar{X}_2)^2}{n_2}} = \sqrt{\frac{23.7143}{7}} = 1.841$

- $S_1^2 = \frac{n_1 S_1^2}{n_1 - 1} = \frac{8(4.186)}{8-1} = 4.784$

- $S_2^2 = \frac{n_2 S_2^2}{n_2 - 1} = \frac{7(3.3893)}{7-1} = 3.95$

i) Null Hypothesis H_0 : $\sigma_1^2 = \sigma_2^2$ i.e. two population have same variance

ii) Alternative Hypothesis H_1 : $\sigma_1^2 > \sigma_2^2$

iii) Level of significance $\alpha = 0.05$

iv) Test statistics : since $S_1^2 > S_2^2$

$$F = \frac{S_1^2}{S_2^2} = \frac{4.784}{3.95} = 1.211$$

v) critical value : $v_1 = n_1 - 1 = 8 - 1 = 7$

$v_2 = n_2 - 1 = 7 - 1 = 6$

$F_{0.05} \; (v_1 = 7, \; v_2 = 6) = 4.20$

vi) Decision :

since $F < F_{0.05}$ the null hypothesis is accepted at 5% level of significance

i.e. Two population have the same variance

(c) Record taken of the number of male and female birth is 830 families having four children as follow

number of male birth	0	1	2	3	4
number of female birth	4	3	2	1	0
number of families	32	178	290	236	94

Test whether the data are consistent with the hypothesis that the Binomial law holds and the chance of male birth is equal to that of female birth, namely $p = q = \frac{1}{2}$ [χ^2 at 5% level of significance for 4 df is 9.49]

$\longrightarrow$

i) Null Hypothesis H_0 : i.e. The data are consistent with the hypothesis of equal probabilities for male and female birth $p = q = 1/2$ i.e it follows Binomial distribution

ii) Alternative Hypothesis H_1 : ~~H~~ The data do not follow Binomial distribution

iii) Level of significance : $\alpha = 0.05$

iv) Test statistics :

Here prob. of getting male birth $p = 1/2$

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

Binomial Distribution :

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

- for calculation of Theoritical frequency

$N(r) = N\, P(X = r)$ here $N = 830$

$n = 4$

$N(0) = 830 \; P(X = 0)$

$= 830 \left[\binom{4}{0} (1/2)^0 (1/2)^4 \right]$

$= 51.875$ $\approx \cancel{52}$

$N(1) = 830 \; P(X = 1)$

$= 830 \times \left[\binom{4}{1} (1/2)^1 (1/2)^3 \right]$

$= 207.5$ $\approx \cancel{208}$

$N(2) = 830 \; P(X = 2)$

$= 830 \times \left[\binom{4}{2} (1/2)^2 (1/2)^2 \right]$

$= 311.25$

$\cancel{\approx 311}$

$N(3) = 830 \; P(X=3)$

$= 830 \left[\binom{4}{3} (1/2)^3 (1/2)^1 \right]$

$= 207.5 \quad \cancel{\approx 208}$

$N(4) = 830 \; P(X=4)$

$= 830 \left[\binom{4}{4} (1/2)^4 (1/2)^0 \right]$

$= 51.875$

$\cancel{\approx 52}$

Observed frequency (f_o)	32	178	290	236	94
Expected frequency (f_e)	51.875	207.5	311.25	207.5	51.875
$(f_o - f_e)^2$	395.02	870.25	451.56	812.25	1774.52
$\frac{(f_o - f_e)^2}{f_e}$	7.615	4.194	1.451	3.914	34.21

$\chi^2 = \frac{\Sigma (f_o - f_e)^2}{f_e} = 51.384$

v) critical value : $\nu = n - 1$

$= 5 - 1$

$= 4$

$\chi^2_{0.05} \; (\nu = 4) = 9.49$

vi) Decision :

since $\chi^2 > \chi^2_{0.05}$ at 5% level of significance

i.e the null Hypothesis is rejected

i.e. The data donot follow binomial Distribution

i.e. The data is not consistent

Q:5

(a) Two sample of size 9 and 8 gives the sum of square of deviation from their respective mean equal 160 inches and 91 inches square ~~so~~ respectively can they be regarded as a drawn from two normal population with same variance?

(F for 8 and 7 d.o.f = 3.73)

→ Here $n_1 = 9$ $n_2 = 8$

$\Sigma(x_1 - \bar{x}_1)^2 = 160$ $\Sigma(x_2 - \bar{x}_2)^2 = 91$

$$S.d. = S_1^2 = \frac{\Sigma(x_1 - \bar{x}_1)^2}{n_1 - 1} = \frac{160}{9-1} = 20$$

$$S_2^2 = \frac{\Sigma(x_2 - \bar{x}_2)^2}{n_2 - 1} = \frac{91}{8-1} = 13$$

i) Null Hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$ i.e. Two population have the same variances

ii) Alternative Hypothesis $H_1 : \sigma_1^2 > \sigma_2^2$

iii) Level of significance $\alpha = 0.05$

iv) Test statistic :

$S_1^2 > S_2^2$

$$F = \frac{S_1^2}{S_2^2} = \frac{20}{13} = 1.538 \approx 1.54$$

$F = 1.54$

v) critical value : $v_1 = n_1 - 1 = 9 - 1 = 8$

$v_2 = n_2 - 1 = 8 - 1 = 7$

~~F.E.O~~ $F_{0.05}(v_1 = 8, v_2 = 7) = 3.73$

vi) Decision : since $F < F_{0.05}$

i.e. The null is accepted at 5% level of significance

∴ The two population have same variance and they do not differ significantly.

(b) A die is thrown 276 times and the result of this throws are given below :

no. appeared on the die	1	2	3	4	5	6
frequency	40	32	29	59	57	59

Test whether the die is biased or not.

[χ^2 at 5% level of significance for 5 d.f. is 11.09]

→ Here $n = 6$

i) Null Hypothesis H_0 : The die is unbiased

ii) Alternative Hypothesis H_1 : The die is biased

iii) Level of significance $\alpha = 0.05$

iv) Test statistics

Here ~~to~~ Expected frequency for each number $fe = \frac{276}{6} = 46$

To find value of χ^2

Observed frequency (fo)	40	32	29	59	57	59
Expected frequency (fe)	46	46	46	46	46	46
$(fo - fe)^2$	36	196	289	169	121	169

$$\chi^2 = \frac{\Sigma (fo - fe)^2}{fe} = \frac{980}{46} = 21.30$$

v) critical value : $\nu = n-1 = 6-1 = 5$

$\chi^2_{0.05} (\nu = 5) = 11.09$

vi) Decision : since $\chi^2 > \chi^2_{0.05}$

The null Hypothesis is rejected

∴ The Die is not unbiased

i.e. The Die is biased

(c) The following figures refers to the observation in live independent samples.

Sample I	25	30	28	34	24	20	13	32	22	38
Sample II	40	34	22	20	31	40	30	23	36	17

Analyse whether the sample have been drawn from the population of equal means

[t at 5% level of significance for 18 d.f is 2.01]

Test whether the means of two population are same at 5% level (t at 0.05 = 2.0739)

→ Here $n_1 = n_2 = n = 10$

	X_1	$(X_1-\bar{X_1})^2$	X_2	$(X_2-\bar{X_2})^2$
	25	2.56	40	114.49
	30	11.56	34	22.09
	28	1.96	22	53.29
	34	54.76	20	86.49
	24	6.76	31	2.89
	20	43.56	40	~~114.49~~ 114.49
	13	184.96	30	0.49
	32	29.16	23	39.69
	22	21.16	36	44.89
	38	129.96	17	151.29
Σ	266	486.4	293	630.08

$$\bar{X_1} = \cancel{266}\ \frac{\Sigma X_{1i}}{n} = \frac{266}{10} = 26.6$$

$$\bar{X_2} = \frac{\Sigma X_{2i}}{n} = \frac{293}{10} = 29.3$$

$\Sigma(X_1-\bar{X_1})^2 = 486.4$ $\Sigma(X_2-\bar{X_2})^2 = 630.08$

i) Null Hypothesis $H_0 : \mu_1 = \mu_2$ i.e. The sample have been drawn from population of equal mean i.e. Their is no significant ~~differnce~~ difference between means

ii) Alternative Hypothesis $H_1 : \mu_1 \neq \mu_2$ (Two tailed test)

iii) Level of significance $\alpha = 0.05$

iv) Test statistics :

$$t = \frac{\bar{X_1} - \bar{X_2}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

here

$$S^2 = \frac{1}{n_1+n_2-2} \left[\Sigma (x_i - \bar{x}_1)^2 + \Sigma (x_2 - \bar{x}_2)^2 \right]$$

$$= \frac{1}{10+10-2} [486.4 + 630.08]$$

$$= 62.0267 \qquad S = 7.875$$

$$t = \frac{26.6 - 29.3}{7.875\sqrt{\frac{1}{10}+\frac{1}{10}}} = -0.76668$$

$$|t| = 0.76668$$

v) critical value :

$$\nu = n_1 + n_2 - 2$$
$$= 10 + 10 - 2$$
$$= 18$$

$$t_{0.05} (\nu = 18) = 2.1$$

vi) decision :

since $|t| < t_{0.05}$

$\therefore$ Null Hypothesis is accepted

$\therefore$ i.e. Their is no significant difference between their means

$\therefore$ i.e Two sample have been drawn from the population of equal means