Model No 5.9: **CHI-SQUARE TEST FOR INDEPENDENT OF ATTRIBUTES**

Problem 28: **The following table gives the classification of 100 workers according to sex and nature of work. Test whether the nature of work is independent of the sex of the worker.**

	Stable	**Unstable**	**Totla**
Males	40	20	60
Females	10	30	40
Total	50	50	100

Solution:

(i) Null Hypothesis (H_0): There is No significant difference b/w observed freq & Expected freq.

(ii) Alternative Hypothesis (H_1): There is Significant difference b/w Observed & Expected frequency

(iii) Level of Significance (α): $\alpha = 0.05$

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

$E(a) = \frac{(a+c)(a+b)}{N} = \frac{60 \times 50}{100} = 30$	$E(b) = \frac{(b+d)(a+b)}{N} = \frac{60 \times 50}{100} = 30$
$E(a) = \frac{(a+c)(c+d)}{N} = \frac{40 \times 50}{100} = 20$	$E(b) = \frac{(b+d)(c+d)}{N} = \frac{40 \times 50}{100} = 20$

Here, No. of Columns $C = 2$

No. of rows $r = 2$

	C↓		
male r→	40	20	60
female	10	30	40
	stable	unstable	100
	50	50	

Ei

30	30
20	20

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
40	30	100	3.3333
20	30	100	3.3333
10	20	100	5
30	20	100	5
			$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 16.6666$

(v) Conclusion: Degrees of freedom $= (n-1)(m-1) = (r-1)(c-1) = (2-1)(2-1) = 1$

Calculated value of $\chi^2_{cal} = 16.6666$

Tabulated value of $\chi^2_{tab} = \chi^2_{0.05}((r-1)(c-1)) = \chi^2_{0.05}(1) = 3.841$

Calculated value of χ^2 Tabulated value of χ^2

$\chi^2_{cal} > \chi^2_{tab}$ Null Hypothesis is Rejected.

Problem 29: **Given the following contingency table for hair colour and eye colour. Find the value of χ^2**

Is there good association between the two?

		Hair colour			
		Fair	**Brown**	**Black**	**Total**
Eye colour	**Blue**	15	5	20	40
	Grey	20	10	20	50
	Brown	25	15	20	60
	Total	60	30	60	150

Solution:

(i) Null Hypothesis (H_0):

(ii)Alternative Hypothesis (H_1):

(iii) Level of Significance (α):

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

$E(a) = \frac{(a+d+g)(a+b+c)}{N} = \frac{40 \times 60}{150}$ 16	$E(b) = \frac{(b+e+h)(a+b+c)}{N} = \frac{40 \times 30}{150}$ 8	$E(b) = \frac{(c+f+i)(a+b+c)}{N} = \frac{40 \times 60}{150} = 16$
$E(c) = \frac{(a+d+g)(d+e+f)}{N} =$ 20 $\frac{50 \times 60}{150}$	$E(b) = \frac{(b+e+h)(d+e+f)}{N} =$ 10 $\frac{50 \times 30}{150}$	$E(b) = \frac{(c+f+i)(d+e+f)}{N} = \frac{50 \times 60}{150} = 20$
$E(c) = \frac{(a+d+g)(g+h+i)}{N} =$ 24 $\frac{60 \times 60}{150}$	$E(b) = \frac{(b+e+h)(g+h+i)}{N} =$ 12 $\frac{60 \times 30}{150}$	$E(b) = \frac{(c+f+i)(g+h+i)}{N} = \frac{60 \times 60}{150} = 24$

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
15	16	1	0.06
5	8	9	1.12
20	16	16	1
20	20	0	0
10	10	0	0
20	20	0	0
25	24	1	0.04
15	12	9	0.75
20	24	16	0.66
			$\chi^2 = \sum \frac{(O-E)^2}{E} = 3.63$

(v) Conclusion: Degrees of freedom $= (n-1)(m-1) = (3-1)(3-1) = 4$

Calculated value of $\chi^2_{cal} = 3.63$

Tabulated value of $\chi^2_{tab} = \chi^2_{0.05}(4) = 9.488$

Calculated value of $\chi^2 <$ Tabulated value of χ^2

Null Hypothesis is Accepted.

Problem 30: **From the following data, find whether there is any significant liking in the habit of taking soft drinks among the categories of employees.**

Soft drinks \ Employees	Clerks	Teachers	Officers	Total
Pepsi	10	25	65	100
Thums Up	15	30	65	110
Fanta	50	60	30	140
Total	75	115	160	350

Solution:

(i) Null Hypothesis (H_0):

(ii) Alternative Hypothesis (H_1):

(iii) Level of Significance (α): $\alpha = 0.05$

(iv) Test Statistic: The test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$

$E(a) = \frac{(a+d+g)(a+b+c)}{N} = \frac{100 \times 75}{350}$	$E(b) = \frac{(b+e+h)(a+b+c)}{N} = \frac{100 \times 115}{350}$	$E(b) = \frac{(c+f+i)(a+b+c)}{N} = \frac{100 \times 160}{350}$
$E(c) = \frac{(a+d+g)(d+e+f)}{N} = \frac{110 \times 75}{350}$	$E(b) = \frac{(b+e+h)(d+e+f)}{N} = \frac{110 \times 115}{350}$	$E(b) = \frac{(c+f+i)(d+e+f)}{N} = \frac{110 \times 160}{350}$
$E(c) = \frac{(a+d+g)(g+h+i)}{N} = \frac{140 \times 75}{350}$	$E(b) = \frac{(b+e+h)(g+h+i)}{N} = \frac{140 \times 115}{350}$	$E(b) = \frac{(c+f+i)(g+h+i)}{N} = \frac{160 \times 160}{350}$

Observed Frequency (O_i)	Expected Frequency (E_i)	$(O_i-E_i)^2$	$\frac{(O_i-E_i)^2}{E_i}$
10	21.4286	130.6129	6.0953
25	32.8571	61.7340	1.8789
65	45.7143	371.9382	8.1361
15	23.5714	73.4689	3.1169
30	36.1429	37.7352	1.0441
65	50.2857	216.5106	4.3056
50	30	400	13.3333
60	46	196	4.2609 18.0625
30	64	1156	
			$\chi^2=\sum\frac{(O_i-E_i)^2}{E_i}$ = 60.2336

(v)

Conclusion: Degrees of freedom $=(n-1)(m-1)=(3-1)(3-1)=4$

Calculated value of χ^2_{cal} = 60.2336

Tabulated value of χ^2_{tab} $\chi^2_{0.05}(4)=9.488$

Calculated value of χ^2 > Tabulated value of χ^2

Null Hypothesis is Rejected

Problem 31: **1000 students at college level were graded according to their I.Q. and the economic conditions of their home. Use χ^2 test to find out whether there is any association between condition at home and I.Q. Use 0.05 L.O.S.**

I.Q. / Economic Condition	High	Low	Total
Rich	460	140	600
Poor	240	160	400
Total	700	300	1000

Solution:

(i) Null Hypothesis (H_0):

(ii) Alternative Hypothesis (H_1):

(iii) Level of Significance (α): $\alpha=0.05$

(iv) Test Statistic: The test statistic $\chi^2=\sum\frac{(O_i-E_i)^2}{E_i}$

$E(a)=\frac{(a+c)(a+b)}{N}=\frac{600\times700}{1000}$ 420	$E(b)=\frac{(b+d)(a+b)}{N}=\frac{600\times300}{1000}$ 180
$E(a)=\frac{(a+c)(c+d)}{N}=\frac{400\times700}{1000}$ 280	$E(b)=\frac{(b+d)(c+d)}{N}=\frac{400\times300}{1000}$ 120

Observed Frequency(O_i)	Expected Frequency(E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
460 140 240 160	420 180 280 120	1600 1600 1600 1600	3.8095 8.8889 5.7143 13.3333
			$\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$ = 31.7460

(v) conclusion: Degrees of freedom $= (n-1)(m-1) = (2-1)(2-1) = 1$

Calculated value of $\chi^2_{cal} = 31.7460$

Tabulated value of $\chi^2_{tab} = \chi^2_{0.05}(1) = 3.841$

Calculated value of χ^2 > Tabulated value of χ^2

Null Hypothesis is Rejected.

Part C:

Analysis of Variance (Anova)

Model No 5.10: One-way Anova

Model No 5.11: Two-way Anova