*Chi-square test :-

Use :-

⇒ Sometimes, data are best collected or conveyed nominally, or categorically. These data are represented by counting the number of times a particular event or condition occurs. In such cases, you may be seeking to determine if a given set of counts, or frequencies, statistically matches some known, or expected, set. Or you may wish to determine if two or more categories are statistically independent. In either case, we can use a non-parametric procedure to analyze nominal data.

*Chi square test for independence :-

⇒ The Chi-square test for independence is used to determine whether there is statistical association between two categorical attributes. The Chi square statistics can be used when two or more categories are involved for two attributes.

General Procedures:

1- H_0 : The two variables of classification are independent.

H_1 : The two variables of classification are not independent (they are associated)

Level of significance:

$$\alpha = 0.05$$

Test Statistic:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - e_{ij})^2}{e_{ij}}$$

O_{ij} = observed frequency.

e_{ij} = expected frequency.

Critical Region:

$\chi^2 \geq \chi^2_{\alpha}(r-1)(c-1)$ reject H_0

Question :

A thousand house holds are taken at random & divided into three groups A, B, C according to the total monthly income test the hypothesis that there is no association b/w total income & television ownership?

Solution :

1- H_0 : The two criteria of classification are independent

H_1 : The two criteria of classification are not independent (they are associated)

2- Level of significance :

$\alpha = 0.05$

3- Test Statistics :

$$\chi^2 = \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{(O_{ij} - e_{ij})^2}{e_{ij}}$$

4- Calculations :

	A	B	C	Total
Colour television	56	51	93	200
Black & white	118	207	375	700
None	26	42	32	100
Total	200	300	500	1000

How to calculate e_{ij}?-

Colour television:	$\frac{200 \times 200}{1000} = 40$.	60	100
Black and white.	140	210	350
None.	20	30	50

O_{ij}	e_{ij}	$(O_{ij} - e_{ij})^2 / e_{ij}$
56	40	6.4
118	140	3.45
26	20	1.8
51	60	1.35
207	210	0.04
42	30	4.8
93	100	0.49
375	350	1.78
32	50	6.48
Total: 1000	1000	26.60.

Critical Region:

$x^2 \geq x^2_{0.05}\,(3-1)(3-1)$

$x^2_{0.05(4)} = 9.49$.

Conclusion:

$X^2 \geq X^2_{0.05(4)}$

$26.6057 \geq 9.49$

Decision: Reject H_0

HOME ASSIGNMENT.

A certain drug is claimed to be effective in curing colds. In an experiment on 164 People with colds, half of them were given the drug & half of them were given suger pills. Test the hypothesis that the drug is no better than sugar pills for curing colds?.

	Helped	Harmed	No effect	Total.
Drug	52	10	20	82
Sugar	44	12	26	82
Total.	96	22	46	164.