UNIT-I

TESTING THE HYPOTHESIS

1.6 Chi – Square test for independence of attributes

An attribute means a quality or characteristic.

Eg: Reading, Drinking, Smoking etc

χ^2test is used to test whether the two attributes are associated or independent.

Let us consider two attributes A and b. A is divided into two classes and B is divided into two classes.

$\boldsymbol{a}$	$\boldsymbol{b}$	$\boldsymbol{a+b}$
c	d	$c+d$
$a+\ c$	$b+d$	N

The expected frequencies are given by

$\boldsymbol{E(a)} = \frac{\boldsymbol{(a+c)(a+b)}}{\boldsymbol{N}}$	$\boldsymbol{E(a)} = \frac{\boldsymbol{(a+c)(a+b)}}{\boldsymbol{N}}$	$\boldsymbol{a} + \boldsymbol{b}$
$\boldsymbol{E(a)} = \frac{\boldsymbol{(a+c)(a+b)}}{\boldsymbol{N}}$	$\boldsymbol{E(a)} = \frac{\boldsymbol{(a+c)(a+b)}}{\boldsymbol{N}}$	$c + d$
$a+\ c$	$b+d$	N

H_0: Attributes are independent.

Degrees of freedom $= (r-1)(s-1)$

Where r – number of rows

S – no of columns

1. An opinion poll was conducted to find the reaction to a proposed civic reform in 100 members of each of the two political parties the information is tabulated below

	Favorable	**Unfavorable**	**Indifferent**
Party A	**40**	**30**	**30**

Party B **42** **28** **30**

Test for independence of reaction with the party affiliations.

Solution:

H_0: Reactions and party affiliations are independent.

The expected frequencies are given by

$\frac{(82)(100)}{200} = 41$	$\frac{(58)(100)}{200} = 29$	$\frac{(60)(100)}{200} = 30$	100
$\frac{(82)(100)}{200} = 4$	$\frac{(58)(100)}{200} = 29$	$\frac{(60)(100)}{200} = 30$	100
82	58	60	200

Observed Frequency	Expected Frequency	(O – E)	$\frac{(o-E)^2}{E}$
40	41	1	0.024
30	29	1	0.024
30	30	0	0
42	41	1	0.034
28	29	-1	0.034
30	30	0	0
200			0.116

Now $\chi^2 = \sum\left[\frac{(o-E)^2}{E}\right] = 0.116$

Number of degrees of freedom $V = (2-1)(3-1) = 2$

Critical value: The tabulated value of χ^2 at 5% for 2 d. f is 5.99

Conclusion:

Since $\chi^2 = 0.116 < 5.99$, then the null hypothesis H_0 is accepted.

i.e., The independence of reactions with the party affiliations may be correct

2. From the following data, test whether there is any association between intelligence andeconomics conditions

Intelligences					
	Excellent	**Good**	**Medium**	**Dull**	**Total**
Economic Conditions	**Good**	**48**	**200**	**150**	**80**	**478**
	Not Good	**52**	**180**	**190**	**100**	**522**
	Total	**100**	**380**	**340**	**180**	**1000**

Solution:

H_0: There is no association between intelligence and economic conditions.

The expected Frequencies are

$\frac{(478)(100)}{1000} = 47.8$	$\frac{(478)(380)}{1000} = 181.64$	$\frac{(478)(340)}{1000} = 162.52$	$\frac{(478)(180)}{1000} = 86.04$	**478**
$\frac{(522)(100)}{1000} = 52.2$	$\frac{(522)(380)}{1000} = 198.36$	$\frac{(522)(340)}{1000} = 177.48$	$\frac{(522)(180)}{1000} = 93.96$	**522**
100	**380**	**340**	**180**	**1000**

Observed Frequency	**Expected Frequency**	**(O – E)**	$\frac{(o-E)^2}{E}$
48	47.8	0.2	0.0008
200	181.64	18.36	0.9645
150	162.52	-12.52	0.0008
80	86.04	-6.04	0.8832

52	52.2	-0.2	1.8558
180	198.36	-18.36	0.4240
190	177.48	12.52	1.6994
100	93.96	6.04	0.3883
1000			6.2168

Now $\chi^2 = \sum\left[\frac{(o-E)^2}{E}\right] = 6.2168$

Number of degrees of freedom $V = (2-1)(4-1) = 3$

Critical value: The tabulated value of χ^2 at 5% for 3 d. f is 7.815

Conclusion:

Since $\chi^2 = 6.2168 < 7.815$, then the null hypothesis H_0 is accepted.

There is no association between intelligence and economics condition

1. **The following data is collected on two characters. Based on this, can you say that there is no relation between smoking and literacy?**

	Smokers	**Non – Smokers**
Literates	**83**	**57**
Illiterates	**45**	**68**

Solution:

H_0: there is no relation between smoking and literacy.

The expected Frequencies are

$\frac{(128)(140)}{253} = 70.83$	$\frac{(125)(140)}{253} = 69.17$	**140**
$\frac{(128)(113)}{253} = 57.16$	$\frac{(125)(113)}{253} = 55.83$	**113**
128	**125**	**253**

Observed Frequency	**Expected Frequency**	**(O – E)**	$\frac{(o-E)^2}{E}$
83	71	12	2.03
57	69	-12	2.09
45	57	-12	2.53
68	56	12	2.57
253			9.22

Now $\chi^2 = \sum\left[\frac{(o-E)^2}{E}\right] = 9.22$

Number of degrees of freedom $V = (2-1)(2-1) = 1$

Critical value: The tabulated value of χ^2 at 5% for 3 d. f is 3.84

Conclusion:

Since $\chi^2 = 9.22 > 3.84$, then the null hypothesis H_0 is rejected.

i.e., there is a relation between smoking and literacy