χ^2 - test to test the Independence of attributes

Let us consider two attributes A and B. A divided into r classes $A_1, A_2, \dots A_r$ and B divided into s classes $B_1, B_2, \dots B_s$. If this is expressed as $r \times s$ matrix, the matrix is called $r \times s$ Contingency table.

<u>Note</u>: For a 2×2 Contingency table

a	b
c	d

$$\chi^2 = \frac{(a+b+c+d)\,(ad-bc)^2}{(a+c)\,(b+d)\,(a+b)\,(c+d)}$$

Observed Freq	Attri B_1	Attri B_2	Total
Attribute A_1	a	b	a+b
Attribute A_2	c	d	c+d
Total	a+c	b+d	a+b+c+d = N

i.e, $$\chi^2 = \frac{N\,(ad-bc)^2}{(a+c)\,(b+d)\,(a+b)\,(c+d)}$$

<u>Calculation of Expected frequency</u>

	Attribute B_1	Attribute B_2
Attribute A_1	$\frac{(a+b)\,(a+c)}{N}$	$\frac{(a+b)\,(b+d)}{N}$
Attribute A_2	$\frac{(a+c)\,(c+d)}{N}$	$\frac{(b+d)\,(c+d)}{N}$

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

Scanned by CamScanner

1. Out of 8000 graduates in a town 800 are females, out of 1600 graduate employees 120 are females. Use χ^2 to determine if any distinction is made in appointment on the basis of sex. Value of χ^2 at 5% Los for one degree of freedom is 3.84.

Solution :

	Male	Female	Total
Graduates in town	7200	800	8000
Graduate employees	1480	120	1600
Total	8680	920	9600

* Ho: There is no significance difference b/w male and female
* H_1 : There is a significance difference b/w male and female.
* Los = 5% , DF = $(r-1)(s-1) = (2-1)(2-1) = 1$
* Table value $\chi^2 = 3.84$
* Test statistic $\chi^2 = \sum \frac{(O-E)^2}{E}$

$$\text{Expected frequency} = \frac{\text{Corresponding row total} \times \text{column total}}{\text{Grand total}}$$

For 7200 $E = \frac{8000 \times 8680}{9600} = 7233.33$

For 800 $E = \frac{8000 \times 920}{9600} = 766.67$

Scanned by CamScanner

For 1480, $E = \frac{(1600)(8680)}{9600} = 1446.67$

120, $E = \frac{(1600)(920)}{9600} = 153.33$

	O	E	(O-E)	$\frac{(O-E)^2}{E} = \chi^2$
	7200	7233.33	-33.33	0.1536
	800	766.67	33.33	1.4490
	1480	1446.67	33.33	0.7679
	120	153.33	-33.33	7.2451
			Total	$9.6156 = \chi^2$

Conclusion: Cal χ^2 $\nless$ Tab. χ^2. So we Reject Ho.

So we ~~accept~~ Ho.

2. 1000 Students at college level were graded according to their IQ and their economic conditions. What conclusion can you draw from the following data.

Economic Condition	IQ Level	
	High	Low
Rich	460	140
Poor	240	160

Soln:

Economic Condition	IQ level High	Low	Total
Rich	460	140	600
Poor	240	160	400
Total	700	300	1000

Observed freq	Expected frequency
460	$\frac{600 \times 700}{1000} = 420$
140	$\frac{600 \times 300}{1000} = 180$
240	$\frac{400 \times 700}{1000} = 280$
160	$\frac{400 \times 300}{1000} = 120$

O	E	$\frac{(O-E)^2}{E}$
460	420	3.809
140	180	8.8889
240	280	5.714
160	120	13.333
	$\sum\left[\frac{(O-E)^2}{E}\right]$ =	31.74

* Ho: There is no significance difference b/w IQ level and economic condition

* H_1 : There is a significance difference b/w IQ level and economic condition

* LOS = 5%   DF $= (r-1)(c-1) = (2-1)(2-1) = 1$

* Table value $\chi^2 = 3.84$

* Test statistic $\chi^2 = \Sigma \left[\frac{(O-E)^2}{E}\right] = 31.74$

* Conclusion : Cal $\chi^2 \nless$ Tab χ^2

We reject Ho.

3. On the bases of information noted below, find out whether the new treatment is comparatively superior to the conventional one.

	Favourable	Non-favourable	
Conventional	40	70	110
New	60	30	90
Total	100	100	200