

Chi-square test

Test for Independence of Attributes

Dr. D S Dhakre

Visva Bharati, Sriniketan, West Bengal, India

Example: A survey of 200 farmers studied whether adoption of improved seed is associated with irrigation availability.

		Improved Seed:		Total
		Yes	No	
Irrigation	Yes	90	30	120
	No	20	60	80
	Total	110	90	200

Test at 5% level of significance whether the Irrigation is associated with Improved.

Solution:

The most suitable measure of contingency between two attributes, called square contingency and denoted by χ^2 is given by

$$\chi^2 = \sum_{i=1}^m \sum_{j=1}^n \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

where in the $m \times n$ contingency table

O_{ij} = observed frequency in $(i, j)^{th}$ cell,

E_{ij} = expected frequency in $(i, j)^{th}$ cell

N = Total frequency

Calculations:

We want to test

H0: Adoption of improved seed and irrigation availability are independent.

Ha: They are associated.

at 5% level of significance

Under null hypothesis, the expected frequencies for each cell are calculated in the table below

		Improved Seed:		
		Yes	No	Total
Irrigation	Yes	$(120 \times 110)/200$ = 66	$(120 \times 90)/200$ = 54	120
	No	$(110 \times 80)/200$ = 44	$(90 \times 80)/200$ = 36	80
	Total	110	90	200

O_{ij}	E_{ij}	$(O_{ij} - E_{ij})$	$(O_{ij} - E_{ij})^2$	$(O_{ij} - E_{ij})^2/E_{ij}$
90	66	24	576	8.7
30	54	-24	576	10.7
20	44	-24	576	13.1
60	36	24	576	16.0
Total				48.5

Chi-square Calculation

$$\chi^2 = \sum_{i=1}^m \sum_{j=1}^n \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 48.5$$

Degree of freedom: $(r - 1)(c - 1) = (2 - 1)(2 - 1) = 1$

Table value of χ^2 at 5% level of significance with 1 df = 3.841

Decision: $48.49 > 3.841$, reject H_0 . Improved seed adoption is significantly associated with irrigation availability.

Decision Rule:

Since $\chi_{cal}^2 > \chi_{tab}^2$, hence H_0 is rejected at 5% level of significance. It means Improved seed adoption is significantly associated with irrigation availability.

Do your self

Example 1: A study was conducted to determine whether the adoption of improved seed varieties is associated with the size of farmers' land holdings.

Land Holding	Adopted Improved Seeds	Did Not Adopt	Total
Small	40	60	100
Large	70	30	100
Total	110	90	200

Example 2: An agricultural researcher wants to know whether the use of fertilizer is associated with crop yield (High or Low). A survey of 200 farmers produced the following data:

Fertilizer Use	High Yield	Low Yield	Total
Used Fertilizer	80	20	100
Did Not Use Fertilizer	50	50	100
Total	130	70	200