

Chi-square test as a test of goodness of fit

Dr Alok Pratap Singh

21/08/2020

The chi-squared test is a widely used non-parametric test in statistical work. Chi-square describes the magnitude of discrepancy between theory and observation.

With the help of Chi-square test, we can find out whether two or more attributes are related or not. As a test of goodness of fit, this test lets us know how appropriately the theoretical distribution fits the empirical distribution.

Following are steps required in testing the goodness of fit.

1 A null hypothesis is established and a significance level is selected for acceptance or rejection of the null hypothesis.

2 A random sample is drawn from a relevant population.

3 A set of expected frequencies is worked out under the assumption that the null hypothesis is true.

4 The observed frequencies are compared with the theoretical or expected frequencies.

5 If the calculated value of Chi-square is less than the table value at a certain level of significance, for certain degrees of freedom, the hypothesis is considered to be good. In such a case the divergence between expected and actual frequency is attributed to fluctuations of sampling.

Here are some numericals on Chi-squared test as a test of goodness of fit:

सांख्यिक रूप से चूने गए 1000 खेतों में उर्वरक के प्रयोग की जांच की गई। निम्नलिखित निष्कर्ष निकाले।

	निज खेत	विद्यार्थी खेत	कुल
उर्वरक का प्रयोग हुआ	416	184	600
उर्वरक का प्रयोग न हुआ	64	336	400
कुल	480	520	1000

क्या आप निष्कर्ष निकाल सकते हैं कि निज खेत में उर्वरक का प्रयोग अधिक होता है? [सार्थकता स्तर 5% है]
 उत्तर: शून्य परिकल्पना - खेत का स्वामित्व और उर्वरक प्रयोग स्वतंत्र हैं।
 प्रस्तावित आकृतियों की सारणी

$\frac{480 \times 600}{1000} = 288$	$\frac{520 \times 600}{1000} = 312$	600
$\frac{480 \times 400}{1000} = 192$	$\frac{520 \times 400}{1000} = 208$	400
480	520	1000

χ^2 (Chi-square) परीक्षण

O (अवलोकित आकृति)	E (प्रत्याशित आकृति)	$(O-E)^2$	$(O-E)^2/E$
416	288	16384	56.889
64	192	16384	85.333
184	312	16384	52.513
336	208	16384	78.769
		$\Sigma (O-E)^2/E = 273.504$	

degrees of freedom = $(r-1)(c-1)$
 $(2-1)(2-1) = 1$

1 d.f. के लिए, $\chi^2_{0.05} = 3.84$

चूंकि आकृतित का स्वरूप (273.504), सारणी के मूल से बहुत अधिक है, अतः शून्य परिकल्पना खारिज की जाती है।
 खेत का स्वामित्व और उर्वरक प्रयोग स्वतंत्र नहीं है।
 खेतों में उर्वरक का प्रयोग अधिक होता है।

Q. 2. निम्न बिजारी के संबंध में पर्यवहार और नई विधि का
 के संख्या 250 लोगों के लिए दिए गए हैं:

विधि	लोगों की संख्या		कुल
	अनुपयोग	उपयोग	
नई	140	30	170
पारंपरिक	60	20	80
कुल	200	50	250

5% पर्यवहार का भी बहाल दिवस नई विधि का
 पारंपरिक विधि से अधिक है।
 उदा: शून्य परिकल्पना: नई और पारंपरिक विधियों के
 बीच अंतर नहीं है।

पर्यवहार आकृतियों की तालिका

$\frac{200 \times 170}{250} = 136$	$\frac{50 \times 170}{250} = 34$	170
$\frac{200 \times 80}{250} = 64$	$\frac{50 \times 80}{250} = 16$	80
200	50	250

χ² गणना:

O	E	(O-E) ²	(O-E) ² /E
140	136	16	16/136 = 0.118
60	64	16	16/64 = 0.250
30	34	16	16/34 = 0.471
20	16	16	16/16 = 1.000

$$\sum (O-E)^2/E = 1.839$$

$$\therefore \chi^2 = \sum (O-E)^2/E = 1.839$$

$$d.f. = (r-1)(c-1) = (2-1)(2-1) = 1$$

5% पर्यवहार का या 1 d.f. के लिए सारणी मूल्य = 3.84

∴ सारणी मूल्य (3.84), अवलंबित मूल्य (1.839) से
 कम है, अतः शून्य परिकल्पना स्वीकार की जाती है।
 नई और पारंपरिक विधि के बीच अंतर स्पष्ट नहीं है।