Chi-square test

Grzegorz Karas

April 19, 2021

1 Definition [1, 2]

The Pearson's χ^2-test is a test used for hypothesis concerning probabilities of multinomial random variables. It tests a null hypothesis stating that the frequency distribution of certain events observed in a sample is consistent with a particular theoretical distribution.

The test statistic is of following kind

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = N \sum_{i=1}^{n} \frac{(O_i/N - p_i)^2}{p_i} \tag{1}$$

where

- χ^2 = Pearson's cumulative test statistic, which asymptotically approaches a χ^2 distribution.
- O_i = number of observation of type i
- N = total number of observations
- $E_i = Np_i$ = the expected (theoretical) count of type i
- n = the number of cells in the table $(rows \cdot columns)$

If the test is fasle, then the appropriate test statistic has approximately a noncentral χ^2 distribution with the same degrees of freedom df and a noncentrality parameter λ, which depends on alternative considered.

2 Test types

Pearson's chi-squared test is used to assess three types of comparison: goodness of fit, homogeneity, and independence.

- A test of **goodness of fit** establishes whether an observed frequency distribution differs from a theoretical distribution.
  χ^2 test statistic with degrees of freedom $df = n - 1$
- A test of **homogeneity** compares the distribution of counts for two or more groups using the same categorical variable.
  χ^2 test statistic with degrees of freedom $df = (rows - 1) \cdot (columns - 1)$
- A test of **independence** assesses whether observations consisting of measures on two variables, expressed in a contingency table, are independent of each other.
  χ^2 test statistic with degrees of freedom $df = (rows - 1) \cdot (columns - 1)$

2.1 Test of goodness of fit

2.2 Test of independence

2.3 Test of homogenity

2.4 Sample size

References

[1] Chi-squared-test `https://en.wikipedia.org/wiki/Chi-squared_test`.

[2] Guenther, W. (1977). *Power and Sample Size for Approximate Chi-Square Tests.* The American Statistician, 31(2), 83-85.