Chi-square test

Grzegorz Karas

August 5, 2021

1 Definition [1, 2]

The χ^2-test is a statistical method to test hypothesis, where random variable follows multinomial distribution. It tests a null hypothesis stating that the frequency distribution of certain events observed in an observer sample is consistent with a particular theoretical distribution.

The most common χ^2-test is a Pearson's chi-square test in which 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 fails, 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 [1]

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

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

2.1 Test of goodness of fit

Let $X = (X_1, ..., X_k)$ be a multinomial ranom variable with parameters $n, p_1, ..., p_k$. Suppose we wish to test

$$H_0: \quad p_i = p_i^0 \qquad i = 1, 2, ..., k \tag{2}$$

against

$$H_a: \text{not all p's are as given by } H_0 \tag{3}$$

where the p_i' are given expected numbers. The value of the chi-square test-statistic is

$$\chi^2_{H_0} = \sum_{i=1}^{k} \frac{\left(x_i - np_i^0\right)^2}{np_i^0} = \sum_{i=1}^{k} \frac{\left(\widehat{p}_i - p_i^0\right)^2}{p_i^0} \tag{4}$$

The chi-square test reject H_0 if

$$\chi^2_{H_0} > \chi^2_{k-1,1-\alpha}, \tag{5}$$

where α is the significance level, and $\chi^2_{k-1,1-\alpha}$ is the quantile of order $1-\alpha$ of the central χ^2 distribution with $k-1$ degrees of freedom. The p-value of the test is

$$p - value = P\left(X > \chi^2_{H_0}\right) \tag{6}$$

To evaluate the power of the test let's precisely define the alternative H_a as follow.

$$H_a: \quad p_i = p_i^a \qquad i = 1, 2, ..., k \tag{7}$$

Thus the power o the test is

$$Power = P^{\lambda}\left(X_a > \chi^2_{k-1,1-\alpha}\right) \tag{8}$$

where X_a is a ranom variable that follows the noncentral χ^2 distribution with the noncentrality parameter

$$\lambda = n \sum_{i=1}^{k} \frac{\left(p_i^a - p_i^0\right)^2}{p_i^0} \tag{9}$$

2.2 Test of independence

Let $X = (X_{ij}) \in \mathbb{R}^{r \times c}$ be a multinomial random variable with parameters n, p_{ij} where $i = 1, 2, ..., r$, $j = 1, 2, ..., c$ and $\sum_{i=1}^{r}\sum_{j=1}^{c} p_{ij} = 1$. Suppopse we wish to test independence

$$H_0: \quad p_{ij} = p_{i\cdot}p_{\cdot j} \qquad i = 1, 2, ..., r \quad j = 1, 2, ..., c \tag{10}$$

against

$$H_a: \text{not all the equatons given under } H_0 \text{ are satisfied} \tag{11}$$

where $p_{i\cdot} = \sum_{j=1}^{c} p_{ij}$ and $p_{\cdot j} = \sum_{i=1}^{r} p_{ij}$. The value of the chi-square test-statistic is

$$\chi^2_{H_0} = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{\left(x_{ij} - x_{i\cdot}x_{\cdot j}/n\right)^2}{x_{i\cdot}x_{\cdot j}/n} \tag{12}$$

where $x_{i\cdot} = \sum_{j=1}^{c} x_{ij}$ and $x_{\cdot j} = \sum_{i=1}^{r} x_{ij}$. We observe that

$$\chi^2_{H_0} = \frac{1}{n}\sum_{i=1}^{r}\sum_{j=1}^{c} \frac{\left(\frac{x_{ij}}{n} - \frac{x_{i\cdot}}{n}\frac{x_{\cdot j}}{n}\right)^2}{\frac{x_{i\cdot}}{n}\frac{x_{\cdot j}}{n}} = \frac{1}{n}\sum_{i=1}^{r}\sum_{j=1}^{c} \frac{\left(\widehat{p_{ij}} - \widehat{p_{i\cdot}}\widehat{p_{\cdot j}}\right)^2}{\widehat{p_{i\cdot}}\widehat{p_{\cdot j}}} \tag{13}$$

The chi-square test reject H_0 if

$$\chi^2_{H_0} > \chi^2_{(r-1)\cdot(c-1),1-\alpha}, \tag{14}$$

To evaluate the power of the test let's precisely define the alternative H_a as follow.

$$H_a: \quad p_{ij} = \underbrace{p_{i\cdot}p_{\cdot j} + \frac{c_{ij}}{\sqrt{n}}}_{p_{ij}^a}, \qquad i = 1, 2, ..., r \quad j = 1, 2, ..., c, \quad where \quad \sum_{i=1}^{r}\sum_{j=1}^{c} c_{ij} = 0, \tag{15}$$

Thus the power of the test is

$$Power = P^{\lambda}\left(X_a > \chi^2_{(r-1)\cdot(c-1),1-\alpha}\right) \tag{16}$$

where X_a is a random variable that follows the noncentral χ^2 distribution with the noncentrality parameter

$$\lambda = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{c_{ij}^2}{p_{i\cdot}p_{\cdot j}} - \sum_{i=1}^{r}\frac{c_{i\cdot}^2}{p_{i\cdot}} - \sum_{j=1}^{c}\frac{c_{\cdot j}^2}{p_{\cdot j}}, \tag{17}$$

where $c_{i\cdot} = \sum_{j=1}^{c} c_{ij}$ and $c_{\cdot j} = \sum_{i=1}^{r} c_{ij}$.

If $\Delta_{ij} = c_{ij}/\sqrt{n}$, then

$$\lambda = \frac{1}{n}\left[\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\Delta_{ij}^2}{p_{i\cdot}p_{\cdot j}} - \sum_{i=1}^{r}\frac{\Delta_{i\cdot}^2}{p_{i\cdot}} - \sum_{j=1}^{c}\frac{\Delta_{\cdot j}^2}{p_{\cdot j}}\right], \tag{18}$$

2.3 Test of homogeneity

Let $X_i = (X_{ij}) \in \mathbb{R}^c$ be a multinomial random variable with parameters n_i, p_{ij} for $i = 1, 2, ..., r$ and $\sum_{j=1}^{c} p_{ij} = 1$. Suppose we wish to test homogeneity

$$H_0: \quad p_{1j} = p_{2j} = \cdots = p_{rj} = p_{\cdot j} \quad j = 1, 2, ..., c \tag{19}$$

against

$$H_a: \text{not all the equatons given under } H_0 \text{ are satisfied} \tag{20}$$

The value of a chi-square test-statistic is

$$\chi^2_{H_0} = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(x_{ij} - x_{i\cdot}x_{\cdot j}/n)^2}{x_{i\cdot}x_{\cdot j}/n} \tag{21}$$

where $x_{i\cdot} = \sum_{j=1}^{c} x_{ij} = n_i$ and $n = \sum_{i=1}^{r} n_i$. The chi-square test reject H_0 if

$$\chi^2_{H_0} > \chi^2_{(r-1)\cdot(c-1),1-\alpha}, \tag{22}$$

To evaluate the power of the test let's precisely define the alternative H_a as follow.

$$H_a: p_{ij} = \underbrace{p_{\cdot j} + \frac{c_{ij}}{\sqrt{n}}}_{p_{ij}^a}, \qquad j = 1, 2, ..., c \quad where \quad \sum_{j=1}^{c} c_{ij} = 0, \tag{23}$$

Thus the power of the test is

$$Power = P^{\lambda}\left(X_a > \chi^2_{(r-1)\cdot(c-1),1-\alpha}\right) \tag{24}$$

where X_a is a random variable that follows the noncentral χ^2 distribution with the noncentrality parameter

$$\lambda = \sum_{j=1}^{c}\frac{1}{p_{\cdot j}}\left[\sum_{i=1}^{r} c_{ij}^2\frac{n_i}{n} - \left(\sum_{i=1}^{r} c_{ij}\frac{n_i}{n}\right)^2\right], \tag{25}$$

If $\Delta_{ij} = c_{ij}/\sqrt{n}$, then

$$\lambda = \frac{1}{n}\sum_{j=1}^{c}\frac{1}{p_{\cdot j}}\left[\sum_{i=1}^{r}\Delta_{ij}^2\frac{n_i}{n} - \left(\sum_{i=1}^{r}\Delta_{ij}\frac{n_i}{n}\right)^2\right] \tag{26}$$

It is worth to observe that $\frac{n_i}{n}$ is the same as $p_{i\cdot}$ and the equation holds following form

$$\lambda = \frac{1}{n} \sum_{j=1}^{c} \frac{1}{p_{\cdot j}} \left[\sum_{i=1}^{r} \Delta_{ij}^2 p_{i\cdot} - \left(\sum_{i=1}^{r} \Delta_{ij} p_{i\cdot} \right)^2 \right] \tag{27}$$

3 Sample size

The sample size required for a test to reach predefined power can be calculated under following assumption

Assumption 1. *The alternative hypothesis is the one given by the observed sample.*

In other words observed estimates construct the alternative hypothesis. The procedure to retrieve the sample size is following:

1. Calculate contingency table in terms of probabilities (p_{ij}^a).
2. For goodness-of-fit set the probabilities, for other test calculate expected values from the contingency table ($\widehat{p_{ij}}$).
3. Calculate deltas Δ_{ij} between contingency table and values from the previous step.
4. Calculate the α-quantile $\chi^2_{df,1-\alpha}$ of a central chi-square distribution.
5. Having defined target power (β) of a test, find the noncentrality parameter (λ) of a noncentral chi-square distribution.

$$\beta = P^{\lambda} \left(X_a > \chi^2_{df,1-\alpha} \right). \tag{28}$$

6. Depending on the type of test calculate the sample size n. Details of the calculation is shown in the next sections.

3.1 Sample size for test of goodness of fit

The p_i^0 from Equation 2 is defined a priori for every i. The p_i^a from Equation 7 is defined a posteriori and is equal to the estimate $\widehat{p_i}$ derived from observed sample for every i. In this case the Equation 9 is following:

$$\lambda = n \sum_{i=1}^{k} \frac{\left(\widehat{p_i} - p_i^0 \right)^2}{p_i^0} \tag{29}$$

So

$$n = \frac{\lambda}{\sum_{i=1}^{k} \frac{\left(\widehat{p_i} - p_i^0 \right)^2}{p_i^0}} \tag{30}$$

3.2 Sample size for test of independence

Having performed steps until 5 the Equation 18 is following

$$\lambda = \frac{1}{n} \left[\sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\Delta_{ij}^2}{\widehat{p_{i\cdot}} \widehat{p_{\cdot j}}} - \sum_{i=1}^{r} \frac{\Delta_{i\cdot}^2}{\widehat{p_{i\cdot}}} - \sum_{j=1}^{c} \frac{\Delta_{\cdot j}^2}{\widehat{p_{\cdot j}}} \right], \tag{31}$$

So

$$n = \frac{1}{\lambda} \left[\sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\Delta_{ij}^2}{\widehat{p_{i\cdot}} \widehat{p_{\cdot j}}} - \sum_{i=1}^{r} \frac{\Delta_{i\cdot}^2}{\widehat{p_{i\cdot}}} - \sum_{j=1}^{c} \frac{\Delta_{\cdot j}^2}{\widehat{p_{\cdot j}}} \right], \tag{32}$$

3.3 Sample size for test of homogeneity

Having performed steps until 5 the Equation 27 is following

$$\lambda = \frac{1}{n} \sum_{j=1}^{c} \frac{1}{\widehat{p_{\cdot j}}} \left[\sum_{i=1}^{r} \Delta_{ij}^{2} p_{i\cdot} - \left(\sum_{i=1}^{r} \Delta_{ij} p_{i\cdot} \right)^{2} \right] \tag{33}$$

So

$$n = \frac{1}{\lambda} \sum_{j=1}^{c} \frac{1}{\widehat{p_{\cdot j}}} \left[\sum_{i=1}^{r} \Delta_{ij}^{2} p_{i\cdot} - \left(\sum_{i=1}^{r} \Delta_{ij} p_{i\cdot} \right)^{2} \right] \tag{34}$$

References

[1] Chi-squared-test `https://en.wikipedia.org/wiki/Pearson%27s_chi-squared_test`.

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