Pearson's Chi-square Test

Chao Cheng

November 3, 2022

There are mainly two types of situations that's suitable for a Pearson's Chi-square test. The first is to test one sample against a given vector, the so-called goodness-of-fit test. And the second is to test the existence of correlation between two samples, the so-called contingency/independence/association test.

1 Effect size index w

The Effect size index w from Chapter 7 in Cohen [2013] is

$$w = \sqrt{\sum_{i=1}^{m} \frac{(P_{1i} - P_{0i})^2}{P_{0i}}}, \tag{1}$$

where

- m is the number of cell.
- P_{0i} is the **propotion** in cell i proposed by the null hypothesis.
- P_{1i} is the **propotion** in cell i proposed by the alternative hypothesis and refects the effect for that cell.

2 Test statistics χ^2

The test statistic is just

$$\chi_T^2 = nw^2 = \sum_{i=1}^{m} \frac{(nP_{1i} - nP_{0i})^2}{nP_{0i}}. \tag{2}$$

3 Goodness of fit test

Let $\boldsymbol{x} \in \mathcal{R}^m$ be a sample from $multinomial(n, \boldsymbol{p})$ where n is the number of trials and $\boldsymbol{p} = (p_1, \cdots, p_m)^T$ and $\sum\limits_{i=1}^{m} p_i = 1$. Here we assume $p_i > 0$ for all i to eliminate some edge cases where some nomial is utterly impossible to happen. Then the probability of any given $\boldsymbol{x} = (x_1, \cdots, x_m)^T$ is

$$P\left(\boldsymbol{x} = (x_1, \cdots, x_m)^T\right) = \prod_{i=1}^{m} p_i^{x_i},$$

where $\sum_{i=1}^{m} x_i = n$. This $\boldsymbol{x}$ can also be seen as the summation of n samples $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ where each $\boldsymbol{x}_i \in \mathcal{R}^m$ follows $multinomial(1, \boldsymbol{p})$. And one and only one entry in each $\boldsymbol{x}_i$ is a single one while others $m-1$ entries all remain zero.

Based on this observed $\boldsymbol{x}$, we want to test its underlying distribution $\boldsymbol{p}$ against a given vector $\boldsymbol{p}_0 = (p_{01}, \cdots, p_{0m})^T$. And from (2) we know that the test statistic is

$$\chi_T^2 = \sum_{i=1}^{m} \frac{(x_i - np_{0i})^2}{np_{0i}}. \tag{3}$$

3.1 Reject rule

Under null hypothesis, this test statistics follows a χ^2 distribution with degree of freedom being $m-1$. Proof for this statement can be found in Chapter 9 Pearson's chi-square test in David R. Hunter's **Notes for a graduate-level course in asymptotics for statisticians** [Hunter, 2014]. And we reject H_0 when this test statistic χ_T^2 is large enough.

3.2 Power analysis

Under alternative hypothesis, i.e. $\boldsymbol{p} = (p_1, \cdots, p_m)^T \neq \boldsymbol{p}_0$. Denote $\boldsymbol{\delta} = \sqrt{n}\,(\boldsymbol{p} - \boldsymbol{p}_0)$ and $\boldsymbol{\Gamma} = \operatorname{diag}(\boldsymbol{p}_0)$. Then the test statistic now follows a **non-central chi-square distribution** with non-central parameter

$$\lambda = \boldsymbol{\delta}^T \boldsymbol{\Gamma}^{-1} \boldsymbol{\delta}.$$

Note: $\lambda = nw^2$, where w is the effect size.

Non-central chi-square distribution: Let $x_1, \cdots, x_n$ be independent normal distribution with means $\mu_1, \cdots, \mu_n$ and unit variance. Then $\sum_{i=1}^{n} x_i^2$ follows a non-central chi-square distribution with non-central parameter being

$$\lambda = \sum_{i=1}^{n} \mu_i^2$$

and degree of freedom being n. And the pdf of $X = \sum x_i$ is given by

$$f(x; n, \lambda) = \exp(-\lambda/2) \sum_{i=0}^{\infty} \frac{(\lambda/2)^i}{i!} f_{n+2i}(x),$$

where $f_n(x)$ stands for the pdf of a ordinary chi-square distribution with n degree of freedom. This result can also be found in Hunter's **Notes for a graduate-level course in asymptotics for statisticians** [Hunter, 2014]. Also Guenther [1977] and Meng and Chapman [1966] offers the same results.

4 Contingency test

The same idea as that in Section 3 for the goodness of fit test except for that $\boldsymbol{p}_0$ is not now given, but rather computed based on **marginal proportion** of the data. So consider a $r \times c$ contingency table in Table 1 and Table 2.

	col_1	$\cdots$	col_c	Total
row_1	x_{11}	$\cdots$	x_{1c}	$x_{1\cdot} = \sum_{j=1}^{c} x_{1j}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
row_r	x_{r1}	$\cdots$	x_{rc}	$x_{r\cdot} = \sum_{j=1}^{c} x_{rj}$
Total	$x_{\cdot 1} = \sum_{i=1}^{r} x_{i1}$	$\cdots$	$x_{\cdot c} = \sum_{i=1}^{r} x_{ic}$	$n = \sum_{i=1}^{r} \sum_{j=1}^{c} x_{ij}$

Table 1: A contingency table, counts in cell

	col_1	$\cdots$	col_c	Total
row_1	$p_{11} = x_{11}/n$	$\cdots$	$p_{1c} = x_{1c}/n$	$p_{1\cdot} = x_{1\cdot}/n$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
row_r	$p_{r1} = x_{r1}/n$	$\cdots$	$p_{rc} = x_{rc}/n$	$p_{r\cdot} = x_{r\cdot}/n$
Total	$p_{\cdot 1} = x_{\cdot 1}/n$	$\cdots$	$p_{\cdot c} = x_{\cdot c}/n$	1

Table 2: A contingency table, proportion in cell

The null hypothesis is that these two types of categories (arranged in row and column, respectively) is independent. Therefore the underlying distribution satisfies

$$p_{ij} = p_{i\cdot} p_{\cdot j}, \quad 1 \le i \le r, \quad 1 \le j \le c. \tag{4}$$

Then the alternative hypothesis is that there exists at least one (i, j) such that (4) does not hold.

4.1 Reject rule

Here the test statistic is

$$\chi_T^2 = n \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(P_{1,ij} - P_{0,ij}\right)^2}{P_{0,ij}} = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(x_{ij} - n p_{i\cdot} p_{\cdot j}\right)^2}{n p_{i\cdot} p_{\cdot j}},$$

where $P_{1,ij}$ is just the observed proportion in cell (i, j) and $P_{0,ij} = p_{i\cdot} p_{\cdot j}$ is the expected proportion computed based on marginal data.

Under null hypothesis, χ_T^2 follows a χ^2 distribution with degree of freedom being $(r-1)(c-1)$. And H_0 is rejected for large value of χ_T^2.

4.2 Power analysis

The same as that in Section 3.2. Just now the $\boldsymbol{p}$ is length $r \times c$ instead of m, and the degree of freedom of the chi-square distribution is $(r-1)(c-1)$.

5 Some conventional assumptions

- Simple random sample: i.i.d sample for each count/trial.
- Sample size(whole table)

- Expected cell count: no zero count. 5 or more in a cell of a 2-by-2 table, and 5 or more in 80% of cells in larger table.

6 Other related tests

- For 2×2 table with small sample size, a Fisher's exact test can be considered.
- For 2×1 table, a binomial test can be considered: Clopper-Pearson's test is an exact one, while the chi-square test or a normal test is a continuous approximation here.

References

Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Routledge, may 2013. doi: 10.4324/9780203771587.

William C. Guenther. Power and sample size for approximate chi-square tests. *The American Statistician*, 31(2):83, may 1977. doi: 10.2307/2683047.

David R. Hunter. Notes for a graduate-level course in asymptotics forstatisticians. June 2014. URL http://personal.psu.edu/drh20/asymp/lectures/asymp.pdf.

Rosa C. Meng and Douglas G. Chapman. The power of chi square tests for contingency tables. *Journal of the American Statistical Association*, 61(316):965–975, dec 1966. doi: 10.1080/01621459.1966.10482187.