Continuity correction of Pearson's chi-square test in 2x2 Contingency Tables: A mini-review on recent development

Nicola Serra [1], Teresa Rea [1], Paola Di Carlo [2], Consolato Sergi [3]

(1) Department of Public Health, University Federico II of Naples, Italy
(2) Department of Sciences for Health Promotion, Mother & Child Care, Univ. of Palermo, Italy
(3) Department of Lab. Medicine and Pathology, Univ. of Alberta, Edmonton, AB, Canada, Stollery Children's Hospital, Univ. of Alberta, Edmonton, AB, Canada

CORRESPONDING AUTHOR: *Nicola Serra Ph.D., Department of Public Health, School of Medicine and Surgery, University Federico II of Naples, Italy. E-mail: nicola.serra@unina.it*

***DOI:* 10.2427/13059**
Accepted on April 16, 2019

ABSTRACT

The Pearson's chi-square test represents a nonparametric test more used in Biomedicine and Social Sciences, but it introduces an error for 2x2 contingency tables, when a discrete probability distribution is approximated with a continuous distribution. The first author to introduce the continuity correction of Pearson's chi-square test has been Yates F. (1934). Unfortunately, Yates's correction may tend to overcorrect of p-value, this can implicate an overly conservative result. Therefore many authors have introduced variants Pearson's chi-square statistic, as alternative continuity correction to Yates's correction. The goal of this paper is to describe the most recent continuity corrections, proposed for Pearson's chi-square test.

Key words: Pearson's x^2 statistic; continuity correction; 2x2 contingency table; Yates's continuity correction, Serra's continuity correction

INTRODUCTION

Pearson's chi-square test or χ^2 test is the nonparametric test commonly used by researchers in Biology, Medicine and Social Sciences. This test is based on the calculation of Pearson's χ^2 statistic, introduced by Pearson K. [1], considering a sample of a population characterized by two o more dichotomous variables. For two dichotomous variables, it is possible to define a 2x2 contingency table, with the frequencies of occurrence of all combinations of their levels, considering a sample size equal to N, as it is shown in Table 1

In a 2x2 contingency table, Pearson's χ^2 statistic is used to test the association between dichotomous variables, for example to individualize a possible association between variables such as sex (Male/Female) and smoke (Yes/No). For this scope Pearson introduce the chi-square statistic to evaluate the discrepancy between observed ($O_{i,j}$) and expected frequencies ($E_{i,j}$), where the observed frequencies are *a, b, c* and d of Tables 1. Instead the expected frequencies are defined for every cell such as:

$$E_{i,j} = \frac{r_i c_j}{N}, \qquad i,j = 1,2$$

where i and j indicate the row and column index respectively. The formula to compute Pearson's χ^2 statistic is described by Pearson K. (1900):

$$\chi_P^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(O_{i,j}-E_{i,j})^2}{E_{i,j}} = \frac{(ad-bc)^2 N}{(a+b)(c+d)(a+c)(b+d)} = \frac{(ad-bc)^2 N}{r_1 r_2 c_1 c_2} \quad [1]$$

where r_1, r_2, c_1 and c_2 i.e. the totals across rows and columns are generally called marginal totals.

Using the χ^2 distribution to interpret Pearson's χ^2 statistic requires one to assume that the discrete probability of observed binomial frequencies of 2x2 contingency table, can be approximated by the continuous χ^2 distribution. This assumption is not entirely correct and introduces some error. To reduce the error in approximation, many authors introduced a continuity correction or variants of Pearson's χ^2 test.

To reduce the error introduced by Pearson's χ^2 statistic, Yates F. [2] suggested a correction for continuity that adjusts the formula for Pearson's χ^2 by subtracting the value 0.5, from the difference between each observed value and its expected value for 2x2 contingency table. This correction reduces the χ^2 value obtained and consequently increases its p-value. The formula to compute Yates's χ^2 statistic in a 2x2 contingency table is:

$$\chi_{Yates}^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(|O_{i,j}-E_{i,j}|-0.5)^2}{E_{i,j}} = \frac{\left(|ad-bc|-\frac{1}{2}N\right)^2 N}{(a+b)(c+d)(a+c)(b+d)} = \frac{\left(|ad-bc|-\frac{1}{2}N\right)^2 N}{r_1 r_2 c_1 c_2} \quad [2]$$

Unfortunately, Yates's correction may tend to overcorrect of p-value; this can implicate an overly conservative result, as reported by several authors [3-7].

The goal of this study is with literature review, to describe the most recent development about the continuity corrections by variants of Pearson's χ^2 test defined for 2x2 contingency tables.

METHODS

In this section we introduce the most recent study about continuity correction of Pearson's χ^2 statistic in 2x2 contingency tables.

Serra's continuity correction

Recently Serra N. [8] introduces a significant minimized of Pearson's χ^2 statistic as a continuity correction of Pearson's χ^2 test, for small samples (sample size ≤ 25). This approach is based on the observation that the denominator r_1 r_2 c_1 c_2 of (1), can be interpreted as a geometric mean. The formula to compute minimize Pearson's χ^2 statistic in a 2x2 contingency table is:

$$\chi_{Serra}^2 = \frac{16}{N^3}(ad-bc) \quad [3]$$

Serra N., showed with a statistical approach, that for small samples (≤25), the minimized Pearson's χ^2 statistic in 2x2 contingency tables, represents a continuty correction for Pearson's χ^2 statistic more effective in comparison to Yates'continuity correction. Particularly in this study the author verify that, the Fisher's exact test [9,10], actually considered the "gold test" used when χ^2 test is not appropriate, i.e. when the sample size is small and the expected values in any of the cells of a 2x2 contingency table are below 5, had performance statistically equal to χ^2_{Serra} test.

Kajita Matchita et al.'s continuity correction

Kajita Matchita et al. [11] proposed a continuity correction to maintain a continuity value to be used when small expected cell frequencies on Pearson's χ^2 test for independence exist in the research data. This correction method is used to control the type I error and obtained using a developed correction in more condition. For this scope the authors used a simulation study. The simulations were performed with Monte Carlo method, to evaluate the performance of their method in comparison to other continuity corrections such as Yates's correction and Williams's correction [12]. It shows an outperformed control of type I error, considering a pattern of data set at a significant level of 0.05 and 0.01, simulated contingency tables between 2x2 and 4x4 (2x2, 2x3, 2x4, 3x3, 3x4

TABLE 1. 2x2 contingency table form.

	Column variable (X)		
Row variable (Y)	*State* 1	*State* 2	*Row totals*
State 1	a	b	$a+b=r_1$
State 2	c	d	$c+d=r_2$
Column totals	$a+c=c_1$	$b+d=c_2$	$\mathbf{N}=a+b+c+d$

and 4x4), a number of small expected cell frequencies up to 30% of the total cell used, a sample size between 5 and 10 times that total cell, and using 10,000 data set simulated by Monte Carlo method for each pattern. The type I error (number rejection of null hypothesis divided by 10,000) was evaluated by Pearson's χ^2 test, i.e. by classical χ^2 test without continuity correction.

In the case of 2x2 contingency tables, where the type I error is greater than the significant level, the χ^2 test equation to be used is as follows:

$$\chi^2_{KM} = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(\left|O_{i,j} - E_{i,j}\right| - C\right)^2}{E_{i,j}} \quad [4]$$

instead, where the type I error is less than the significant level, the χ^2 test equation is

$$\chi^2_{KM} = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(\left|O_{i,j} - E_{i,j}\right| + C\right)^2}{E_{i,j}} \quad [5]$$

where $O_{i,j}$ and $E_{i,j}$ represent the observed and expected frequencies respectively, instead C is the developed correction value. It was computed in two cases as follows, if the type I error is higher than the significant level, the authors try to replace the value C into equation (4) start from 0.01, 0.02, 0.03, ..., . If the type I error is less than the significant level, they try to replace the value C into equation (5) start from 0.01, 0.02 , 0.03 ..., . After they replaced value C and computed type I error then to compared with significant level. Developed correction value (C) is the value which gets very similar values between type I error and significant level.

CONCLUSION

In this paper we described the most recent studies of continuity correction of Pearson's χ^2 test. Since the first continuity correction proposed by Yates (1934), produced an overcorrection of the p-value, many authors are discouraging its use. Instead other authors [13-18], have followed Yates (1934) in claiming that the use of Pearson's χ^2 in the case of 2x2 contingency tables tends to generate too many type I errors, especially with small samples, therefore they defined different continuity corrections of Pearson's χ^2 statistic, to reduce the type I error, and simultaneously to reduce the type II error that Yates's correction introduces

Unfortunately, the study of continuity correction of Pearson's χ^2 statistic is very limited in the recent statistical literature, only two recent studies are dedicated at this problem (Serra N., 2018 and Kajita Matchita et al., 2018), showing of the variants of χ^2 statistic as continuity correction of Pearson's χ^2 test.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not for profit sectors.

Competing interests statement

There are no competing interests for this study.

References

1. Pearson K. (1900), On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Philosophical Magazine Series 5; 50(302):157–175.
2. Yates, F. (1934). Contingency tables involving small numbers and the χ^2 test. Supplement to the Journal of the Royal Statistical Society, 1(2), 217-235.
3. Camilli, G., & Hopkins, K. D. (1978). Applicability of chi-square to 2x2 contingency tables with small expected cell frequencies. Psychological Bulletin, 85(1), 163.
4. Campbell, I. (2007). Chi-squared and Fisher–Irwin tests of two-by-two tables with small sample recommendations. Statistics in Medicine, 26(19), 3661-3675.
5. Haber, M. (1982). The continuity correction and statistical testing. International Statistical Review/Revue Internationale de Statistique, 135-144.
6. Richardson, J. T. (1990). Variants of chi-square for 2×2 contingency tables. British Journal of Mathematical and Statistical Psychology, 43(2), 309-326.
7. Richardson, J. T. (2011). The analysis of 2×2 contingency tables—Yet again. Statistics in Medicine, 30(8), 890-890.
8. Serra, N. (2018). A significant minimization of Pearson's c2 statistics in 2x2 contingency tables: preliminary results for small samples. Epidemiology, Biostatistics and Public Health, 15(3).
9. Agresti, A. (2001). Exact inference for categorical data: recent advances and continuing controversies. Statistics in medicine, 20(17-18), 2709-2722.
10. Fisher, R.A. (1934), Statistical Methods for Research Workers. Chapter 12. 5th Ed., Oliver & Boyd.
11. Matchima, K., Vongprasert, J., & Chutiman, N. (2018). The Development of a Correction Method for Ensuring a Continuity Value of The Chi-square Test with a Small Expected Cell Frequency. Naresuan University Journal: Science and Technology (NUJST), 26(1), 98-105.
12. Mcdonald, J.H. (2014), Handbook of Biological Statistics. Maryland: Sparky House Publishing.
13. Cochran WG. (1954), Some methods for strengthening the common χ^2 tests. Biometrics; 10(4):417–451.
14. Cox, D.R. (1970). The continuity correction. Biometrika 57: 217-219.
15. Feller, W. (1968). An Introduction to Probability Theory and Its Applications. Volume I, 3rd ed. John Wiley & Sons, Inc. New York.
16. Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute, 22(4), 719-748.
17. Maxwell, E. A. (1976). Analysis of contingency tables and further

reasons for not using Yates correction in 2× 2 tables. The Canadian Journal of Statistics/La Revue Canadienne de Statistique, 277-290.
18. Upton, G. J. G. (1982). A comparison of alternative tests for the 2× 2 comparative trial. Journal of the Royal Statistical Society. Series A (General), 86-105.

*