Academic Journal of Applied Mathematical Sciences
ISSN(e): 2415-2188, ISSN(p): 2415-5225
Vol. 5, Issue. 1, pp: 1-6, 2019
URL: https://arpgweb.com/journal/journal/17
DOI: https://doi.org/10.32861/ajams.51.1.6

Academic Research Publishing Group

Original Research Open Access

Alternative Methods of Solving Biasedness in Chi – Square Contingency Table

Okeke Charles C.
Department of Statistics, Nnamdi Azikiwe University, Awka

Abstract

Some methods of solving the biasedness in Chi–square Contingency table statistic were considered. Phi Coefficient, Contingency Coefficient and Cramer's V tools were employed to solve the biasedness in the use of Chi–square test. Our results show that any of Phi coefficient, Contingency coefficient or Cramer's V can be used to describe the association between two variables if the data matrix is 2 x 2. Contingency Coefficient was recommended as a good statistic when the matrix dimension is the same while the Cramer's V is most adequate when the data matrix differs.
Keywords: Biasedness; Contingency table; Phi coefficient; Cramer's V.

1. Introduction

Chi-square statistics are commonly used for tests of fit of measurement models. Chi-square is also sensitive to sample size, which is why several approaches to handle large samples in test of fit analysis have been developed [1]. One method to handle the sample size problem may be to adjust the sample size in the analysis of fit or to adopt a random sample method.

Categorical data collected based on complex sample design is not proper for the standard Pearson multinomial-based Chi-squared test because the observations are not independent and identically distributed [2]. He investigates the effects of bias of point estimator of population proportion and its variance estimator to the Pearson Chi-Squared test statistics when the sample is collected based on complex sampling scheme.

Chi-squared test is also influenced by the number of observations obtained. This test assesses departure from independence with values of χ^2 close to zero indicating independence and large values reflecting lack of independence of the two variables. It becomes better to consider measure of association that does not depend on the number of observations. One problem with the test is that it does not indicate the nature of relationship. It is not possible to determine the extent to which one variable changes as the values of the other variable changes. Another setback of the use of Chi-squared test is that the size of the Chi-square statistic does not provide a reliable guide to the strength of the statistical relationship between the two variables. This is evident when two different cross classification tables have the same sample size, the two variables having larger chi-square statistic are more related than are the variables in the table with less chi-square statistic. A misleading result also exists when the sample sizes differ for two tables. Finally, the chi-square value may change depending on the cell numbers.

Karl Pearson suggested the simplest of the measures of association based on χ^2 called Phi Coefficient [3]. It has two major drawbacks as a measure of association. Firstly, it has no simple interpretation when compared to different sets of data. The second disadvantage is that it is not guaranteed to lie between 0 and 1.

The correlation between two items is the usual product-moment correlation between two variables, where variables are restricted to the integers 1 or 0. This statistic is the phi coefficient and is applicable to 2 x 2 tables only. When investigators speak of the correlation between dichotomously scored test items, the reference is the phi coefficient [4].

Another chi squared based measure of association that also adjusts for different sample sizes is the Contingency coefficient (C). The minimum value of C is 0 while the maximum values of the contingency coefficient depends on the number of categories of the variables. Contingency coefficient is not directly comparable unless calculated on tables containing the same number of rows and columns [4].

Albrecht [5], described the contingency table as the relationship between two categorical variables. Contingency table helps to determine the effectiveness of a system under study or if the effectiveness of the system is based on certain Profile variables.

Agresti [6], said that association refers to coefficients which gauge the strength of a relationship. Coefficients in this section are designed for use with nominal data. Phi and Cramer's V are based on adjusting chi – square significance to factor out sample size. These measures do not lend themselves to easy interpretation. Phi and Cramer's V vary between 0 and 1.

Cramer's V (V) is another measure of association between two nominal variables with a range of 0 and 1 inclusive. V was published by Cramer [7] and is also based on Pearson's chi-squared statistic. Cramer's V is a symmetric measure which is independent of the placement of the variables in either the rows or columns. In a case of 2 x 2 contingency table, Cramer' V is equal to phi coefficient. It is also used when the number of rows and columns are unequal.

Cramer's V may also be applied to goodness of fit chi-squared models when there is a 1 x k table. In this case k is taken as the number of optional outcomes and it functions as a measure of tendency towards a single outcome [8].

Campbell [9], compared Chi-square analyses of 2 × 2 tables for many different sample sizes and designs and found that a statistic suggested by Karl Pearson's son [10] called the N - 1 chi-square test provided the best estimates. He added as long as the expected frequency is at least 1, this adjusted Chi-square probably the "Pearson correction" referred to in the question at the top of this column provided the most accurate estimates of Type I error levels. However, for expected frequencies below 1, he found that Fisher's exact test performed better.

Mutai [11], contended that goodness of fit is way of comparing empirically derived data (expected as frequencies) with theoretically expected results. In other words, a research situation may occur in which the experimenter is interested in similarity (or "goodness of fit") between the distribution of a sample of observations, and the distribution of cases that previous research or theory would suggest [11].

Howell [12], proposed that the chi-square test is hypothesis testing technique that produces statistics that is approximately distributed as the chi square distribution.

Bartlett [13] quoted from a letter of E.S. Pearson dated 30 March 1979: I knew long ago that Karl Pearson used the 'correct' degrees of freedom for (a) difference between two samples and (b) multiple contingency tables. But he could not see that χ^2 in curve fitting should be got asymptotically into the same category.

Ingersoll [14], noted that a 2 x 2 contingency table is used to conceptualize, organize and report data. The null hypothesis tested with a chi-square test based on a 2 x 2 contingency table is considered as test of independence. He also added a well-established fact that the expected frequency should be at least five to enable the application of the chi-square meaningful otherwise other techniques are applied.

Rosenberg [15], proposed that phi is the square root of Chi – square divided by n, the sample size: phi = SQRT (χ^2 / n). When computing phi, note that Yates correction to Chi – square is not used. phi measures the strength of the relationship defined as the number of cases on one diagonal minus the number on the other diagonal, adjusting for the marginal distribution of the variables. He observed that some computer packages, such as systat, use special formulas for phi in 2 – by - 2 tables so that phi varies from - 1 to + 1, allowing it to indicate.

Sokal and Rohlf [16], observed the effect of Yates's correction is to prevent over-estimation of statistical significance for small data. The sample size must be large enough to fairly represent the population from which it is drawn. At least 20 observations should be used, with at least five members in every individual category. They added that unfortunately, Yates's correction may tend to overcorrect. This can result in an overly conservative result that fails to reject the null hypothesis when it should (a type II error).

Daniel [17], showed that the Chi-square issues expected frequency cell size, the greater the numerical difference between observed and expected frequencies within cells, the more likely is a statistically significant χ^2. Cells with estimated frequencies (<5) may yield an inflated χ^2. There is considerable debate concerning small cell expected frequencies.

Oliver and Bell [18], suggested odds ratios of 2, 3, and 4 better correspond with small, medium, and large effects, for other related measures of effect size for categorical outcomes.

Overall [19], examined the effect of low expected frequencies in one row or column of 2 x 2 design on the power of the chi-square statistic. This most often results from the analysis of infrequently occurring events. Setting (1 — a) = .70 as a minimally acceptable level. He concluded that when expected values are quite low, the power of the Chi-square test drops to a level that produces a statistic that, in his view, is almost useless.

Wallis [20], suggested that it is possible to employ statistically sound methods for comparing different sizes of effect by estimating a Gaussian confidence interval or by comparing contingency tables for the difference of differences. He also discussed goodness of fit measures of association which measure the degree to which one categorical distribution correlates with another.

Wallis [21], demonstrated that Yates 'continuity corrected' χ^2 obtains a closer approximation to the Binomial distribution than standard χ^2, and notes that Newcombe's continuity corrected interval also mirrors this adjusted Gaussian.

2. Methodology

In order to detect and solve the biasedness in the use of chi squared statistic, phi coefficient, contingency coefficient and Crammer's V statistics will be employed.

2.1. Phi Coefficient (ϕ)

In measure of association, Phi is used to adjust the Chi square statistic by the sample size. Phi is defined as

$$\phi = \sqrt{\frac{\chi^2}{n}} \qquad (2.1.1)$$

χ^2 is the chi-square statistics and n is the sample size.

To determine the value of ϕ, the value of the χ^2 is calculated from the contingency table.

2.1.2. Contingency Coefficient (C)

Contingency coefficient is another measure of association that adjusts for sample sizes and is given by

$$C = \sqrt{\frac{\chi^2}{n+\chi^2}} \qquad (2.1.2)$$

When C = 0, the interpretation is that there is no relationship between the two variables. However, the value of C cannot exceed 1.

2.1.3. Cramer's V (V)

The last statistic commonly used by researchers in measuring association between nominal variable is the Cramer's V statistics and is given by

$$V = \sqrt{\frac{\phi^2}{k}} = \sqrt{\frac{\chi^2}{nk}} \qquad (2.1.3)$$

Where $k =$ Minimum (r-1)(c-1)

Where r and c are the number of rows and columns respectively.

If the dimension of the contingency table is 2 x 2, then r – 1 = c – 1 = k.

$$V = \sqrt{\frac{\phi^2}{k}} = \sqrt{\frac{\phi^2}{1}} = \phi$$

Therefore

3. Illustrative Example

In other to solve the biasedness in chi-square test, the following illustrative examples will be used.

3.1. 2 x 2Contingency Table

Of a group of patients who complained that they did not sleep well, some were given sleeping pills while others were given sugar pills (although they all thought they were getting sleeping pills). They were later asked whether the pills help them or not. The results of their response are shown 3.1.1.

Table-3.1.1. 2 x2 Contingency Table of sleeping pills and mode of sleep

	Slept well	Did not sleep well
Took sleeping pills	44	10
Took sugar pills	81	35

Source: [22]

From Table 3.1.1 above, the Chi-square calculated is 2.571 while the p-value is 0.109. Since the p-value of 0.109

> 0.05, there are no association between the sleeping pills and mode of sleep.

1,000 students at college level were graded according to their I. Q. and the economic conditions of their homes.

Table-3.1.2. 2 x2 Contingency Table of economic conditions and I.Q

	High	Low
Rich	460	140
Poor	240	160

Source: Statistical methods, forty second edition (2012) by S.P. Gupta, Sultan Chand and Sons)

From Table 3.1.2 above, the Chi-square calculated is 31.746 while the p-value is 0.000. Since the p-value of 0.000 < 0.05, there is an association between economic conditions and I.Q.

Table-3.1.3. 2 x2 Contingency Table of sleeping pills and mode of sleep multiplied by 4

	Slept well	Did not sleep well
Took sleeping pills	176	40
Took sugar pills	324	140

Source: [22]

From Table 3.1.3 above, the Chi-square calculated is 10.284 while the p-value is 0.001. Since the p-value of 0.001 < 0.05, there is an association between the sleeping pills and mode of sleep.

Table-3.1.4. 2 x2 Contingency Table of sleeping pills and mode of sleep divided by 5.

	Slept well	Did not sleep well
Took sleeping pills	9	2
Took sugar pills	16	7

Source: [22]

From Table 3.1.4 above, the Chi-square calculated is 0.574 while the p-value is 0.449. Since the p-value of 0.449 > 0.05, there is no association between the sleeping pills and mode of sleep.

Table-3.2. Summary of a 2 x 2 table

	Strong n = 1000			**Weak n = 170**		**Weak n = 680**		**Weak n = 34**	
	High	**Low**		**Slept well**	**Did not sleep well**	**SW**	**DNSW**	**SW**	**DNSW**
Rich	460	140	Took sleeping pills	44	10	176	40	9	2
Poor	240	160	Took sugar pills	81	35	324	140	16	7
χ^2	31.746			2.571		10.284		0.574	
N	1000			170		680		34	
ϕ^2	0.032			0.015		0.015		0.017	
ϕ	0.178			0.122		0.122		0.130	
C	0.175			0.122		0.122		0.130	
V	0.178			0.122		0.122		0.130	

Source: Statistical methods, forty second edition (2012) by S.P. Gupta, Sultan Chand and Sons / [22]

3.3. 3 x 3 Contingency Table

Three samples are taken comprising 120 doctors, 150 advocates and 130 university teachers. Each person chosen is asked to select one of the three categories that best represents his feeling toward a certain national policy. The three categories are in favour of policy (F), against the policy (A), and indifferent toward the policy (I). The results of the interviews are given below:

Reaction

Occupation	**F**	**A**	I	**Total**
Doctors	80	30	10	120
Advocates	70	40	40	150
University Teachers	50	50	30	130
Total	200	120	80	400

Source: Statistical methods, forty second edition (2012) by S.P. Gupta, Sultan Chand and Sons

Table-3.4. Summary of a 3 x 3 table

Reaction									
	Strong n = 200			**n = 528**			**n = 800**		
Occupation	F	A	I	F	A	I	F	A	I
Doctors	40	15	5	80	30	10	160	60	20
Advocates	35	20	20	70	40	40	140	80	80
University Teachers	25	25	15	50	50	30	100	100	60
χ^2	13.618			27.235			54.470		
φ	0.261			0.261			0.261		
C	0.252			0.252			0.252		
V	0.261			0.261			0.261		

Source: [22]

3.5. 3 x 2 Contingency Table

The following table gives the number of good and bad parts produced by each of three shifts in a factory.

Table-3.1.1. Association between the shift and the quality of parts produced.

Shift	**Good**	**Bad**	**Total**
Day	900	130	1030
Evening	700	170	870
Night	400	200	600
	2000	500	2500

Source: Statistical methods, forty second edition (2012) by S.P. Gupta, Sultan Chand and Sons

Table-3.6. Summary of a 3 x 2 table

Shift	Good	Bad	Good	Bad	Good	Bad
Day	300	43.3	900	130	1800	260
Evening	233.3	56.7	700	170	1400	340
Night	133.3	66.7	400	200	800	400
χ^2	34.681		101.830		203.660	
φ	0.205		0.202		0.202	
C	0.191		0.197		0.197	
V	0.205		0.202		0.202	
N	833		2500		5000	

Source: Statistical methods, forty second edition (2012) by S.P. Gupta, Sultan Chand and Sons

4. Discussions and Conclusions

In order to correct the biasedness in chi-squares statistic, phi coefficient, contingency coefficient and crammer's V statistics were administered on 2 x 2, 3 x 3, and 3 x 2 contingency tables.

In tables 3.1.1 and 3.1.4, the values of ϕ, C and V are the same while 3.1.2 and 3.1.3, the values of ϕ, C and V are not the same. Also, tables 3.1.3 and 3.1.4 which is based on the weak relationship of Table 1 but the sample size increased from n = 170 to n = 680. Each of the observed numbers of cases in the cell of Table 3.1.1 are multiplied by 4, in other to preserve the nature of the relationship and table 3.1.4 which is based on the weak relationship of Table 1 but the sample size increased from n = 170 to n = 34. Each of the observed numbers of cases in the cell of Table 3.1.1 are divided by 5, in other to preserve the nature of the relationship. However, in table 3.4, is analyzed in three ways (i) the left panel was divided by 2 (ii) the middle panel is the same with Table 3.3 with same size (iii) the right panel is multiplied by two. From the summary of the researcher clearly shows the biasedness of Chi square association, the Chi square values increase as the sample increases. The table is arranged in ascending order with different Chi square values of test of association but was corrected with Phi Coefficient (φ) , Contingency Coefficient (C) and Cramer's V (V). Phi Coefficient (φ) has the test of association as 0.261 for all the three tables, Contingency Coefficient (C) as 0.252 for the three tables and Cramer's V as 0.261 for the three tables and the value of Cramer's V are same from the value of Phi Coefficient. This tells us that there is an association between the variables. The 3 x 2 table above has its Chi square association differ in the three tables. The first is 34.681 the first table that was reduced by 3, second as 101.830 the main table and 203.660 in the last table that was multiplied by 2. This shows that no matter the dimension of a table the Chi square is always affected by the sample size. Corrected by Phi Coefficient as 0.205, 0.202 and 0.202 for the three tables, Contingency Coefficient as 0.191, 0.197 and 0.197 & Cramer's V as 0.205,0.202 and 0.202. There all shows a weak association between the shift and the quality of parts produced.

The appropriate measure of correlation when working with nominal data is determined by the structure of the data matrix and the characteristics of the data. If the data matrix is 2 x 2, any of the phi coefficient, contingency coefficient or Cramer's V can be used. However, phi statistics is highly recommended. In the same way, if the dimension of the data matrix is the same, the Contingency Coefficient is recommended. Finally, Cramer's V is recommended if the data matrix is unequal.

References

[1] Bergh, D., 2015. "Chi-squared test of fit and sample size-a comparison between a random sample approach and a chi-square value adjustment method." *J. Appl Meas*, vol. 16, pp. 204-217.

[2] Sunyeong, H., 2012. "Effect of bias on the pearson chi-squared test for two population homogeneity test." *Journal of the Chosun Natural Science* vol. 5, pp. 241-245.

[3] Leach Chris, 1978. *Introduction to statistics-a nonparametric approach for the social sciences*. Newcastle upon Tyne, p. 339.

[4] Ferguson, G. A., 1981. *Statistical analysis in psychology and education*. McGraw-Hill Kogakusha, Ltd.: 549.

[5] Albrecht, C., 2013. "Contingency table and the chi square statistics, Interpreting computer printouts and constructing tables." Available: http://extension.usu.edu/evaluation/files/uploads/Start%20Your%20Engine/Study%20the%20Route/Analyze%20the%20Data/Interpreting_Chi_Square_Printouts.pdf

[6] Agresti, A., 2006. *Introduction to categorical data analysis*. New York: John Wiley and Sons.

[7] Cramer, H., 1946. *Mathematical methods of statistics*. Princeton: Princeton University Press. p. 282.

[8] Sheskin, D. J., 1997. *Handbook of parametric and nonparametric statistical procedures*. Boca Raton, FI: CRC Press.

[9] Campbell, I., 2007. "Chi-squared and Fisher-Irwin tests of two-by-two tables with small sample recommendations." *Statist. Med.*, vol. 26, pp. 3661-3675.

[10] Pearson, E. S., 1947. "The choice of statistical tests illustrated on the interpretation of data classed in a 2 x 2 table." *Biometrika*, vol. 34, pp. 139-167.

[11] Mutai, B. K., 2000. *How to write quality research proposal, A complete and simplified recipe*. New Delhi: Thelley Publications.

[12] Howell, D. C., 2012. "Chi-square test-analysis of contingency tables." Available: http://www.uvm.edu/~dhowell/methods8/supplement

[13] Bartlett, M. S., 1981. "Egon sharpe pearson, 1895-1980." *Biometrika*, vol. 68, pp. 1-12.

[14] Ingersoll, G. M., 2010. "Analysis of 2x2 contingency tables in educational research and evaluation." *International Journal for Research in Education*, vol. 27, pp. 1-14. Available: http://www.cedu.uaeu.ac.ae/journal/issue27/ch5_27en.pdf

[15] Rosenberg, M., 2000. *The logic of survey analysis*. New York: Basic Books.

[16] Sokal, R. R. and Rohlf, F. J., 1981. *Biometry, The principles and practice of statistics in biological research*. Oxford: W.H. Freeman.

[17] Daniel, W. W., 1990. *Applied nonparametric statistics*. 2nd ed. Boston: PWS Kent Publishing Company.

[18] Oliver, J. and Bell, M. L., 2013. "Effect sizes for 2 x 2 contingency tables." *PLOS One*, vol. 8, p. e5877.

[19] Overall, J. E., 1980. "Power of chi-square tests for 2 X 2 contingency tables with small expected frequencies." *Psychological Bulletin*, vol. 87, pp. 132-135.

[20] Wallis, S. A., 2011. *Comparing tests for separability. Interval estimation for the difference between a pair of differences between two proportions*. London: Survey of English Usage, UCL.

[21] Wallis, S. A., 2013. "Binomial confidence intervals and contingency tests." *Journal of Quantitative Linguistics*, vol. 20, pp. 178-208.

[22] Spiegel, M. R. and Stephens, L. J., 2008. *Schaum's outline statistics*. 4th ed. The McGraw Hills, pp. 313-314.