SIMPLIFIED CHI-SQUARE STATISTIC (C-SQUARE)

F.E. Amoyedo[1], W. B. Yahya[2] and A.O. Adeoye[3]

[1,2] Department of Statistics, University of Ilorin, Ilorin, Nigeria
[3] Department of Mathematics and Statistics, Federal Polytechnic Offa

Corresponding Author: Amoyedo F. E., amoyedofemi@gmail.com

ABSTRACT: *This research "simplified chi – square (C-square)" is carried out to remove the difficulties in the conventional chi-square. In this paper, a chi-square statistic is simplified to obtain a statistic called C^2. C-square is the sum of square of the co-factor divided by marginal totals multiply by ground total of any given categorical data in which the estimate obtained from the statistic is almost the same with the estimate from chi-square. C^2or C-square can be derived from chi-square and also be obtained through co-factor method This research work is carry out in order to check the cumbersome and difficulties in chi-square manual computation and minimize the time spending in it computation and also reduce the work load, data entering and time taking for R-programming to execute the assignment. When C^2or C-square is used with several examples we discovered that it is faster and the result is approximately the same as the result obtain in R programming and conventional chi – square.*

KEYWORDS: *C-square,chi-square, cofactor, statistic, R-programming.*

1. INTRODUCTION

In statistics, a categorical variable is a variable that can take on one of a limited, and usually fixed number of possible values, assigning each individual or other unit of observation to a particular group or nominal category on the basis of some qualitative property. A categorical variable that can take on exactly two values is termed a binary variable or dichotomous variable; an important special case is the Bernoulli variable. Categorical variables with more than two possible values are called polytomous variables; categorical variables are often assumed to be polytomous unless otherwise specified. [6]

2. METHODOLOGY

2.1 Chi-square method

A **chi square** (χ^2) statistic is a test that measures how expectations compare to actual observed data (or model results). The data used in calculating a **chi square** statistic must be random, raw, mutually exclusive, drawn from independent variables, and drawn from a large enough sample [7].

For any (n x n) contingency table, it can be written as $X^2 = \sum_{i=1}^{r}\sum_{i=1}^{c}\frac{(n_{ij}-e_{ij})^2}{e_{ij}}$

The Proposed C-square is the sum of square of the co-factor divided by marginal totals multiply by ground total of any given categorical

data = $C^2 = \sum_{i=1}^{r}\sum_{i=1}^{c}\frac{(C_{ij})^2}{n\, n_{i.}n_{.j}}$

where $C_{ij} = \det\begin{pmatrix} n_{ij} & n_{i.} \\ n_{.j} & n \end{pmatrix}$ I, j= 1,2,3

Proof

Table 1: showing (3 x 3) contingency table in a categorical form

	B			Total
A	n_{11}	n_{12}	n_{13}	$n_{1.}$
	n_{21}	n_{22}	n_{23}	$n_{2.}$
	n_{31}	n_{32}	n_{33}	$n_{3.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.3}$	n

where

$n_{11} + n_{12} + n_{13} = n_{1.}$

$n_{21} + n_{22} + n_{23} = n_{2.}$

$n_{31} + n_{32} + n_{33} = n_{3.}$

$n_{11} + n_{21} + n_{31} = n_{.1}$

$n_{12} + n_{22} + n_{32} = n_{.2}$

$n_{13} + n_{23} + n_{33} = n_{.3}$

$n_{11} + n_{12} + n_{13} + n_{21} + n_{22} + n_{23} + n_{31} + n_{32} + n_{33} + n_{11} + n_{21} + n_{31} + n_{12} + n_{22} + n_{32} + n_{13} + n_{23} + n_{33} = n$

Obtaining the expected Table 1 above

$e_{11} = \frac{n_{1.}\, X\, n_{.1}}{n}$ $e_{12} = \frac{n_{1.}\, X\, n_{.2}}{n}$ $e_{13} = \frac{n_{1.}\, X\, n_{.3}}{n}$

$e_{21} = \frac{n_{2.}\, X\, n_{.1}}{n}$ $e_{22} = \frac{n_{2.}\, X\, n_{.2}}{n}$ $e_{23} = \frac{n_{2.}\, X\, n_{.3}}{n}$

$e_{31} = \frac{n_{3.} X\, n_{.1}}{n}$ $\quad e_{32} = \frac{n_{3.} X\, n_{.2}}{n}$ $\quad e_{33} = \frac{n_{3.} X\, n_{.3}}{n}$

Substitute the expectation into the chi-square formulae

$$X_{11}^2 = \frac{n_{11} - (\frac{n_{1.}n_{.1}}{n})}{\frac{n_{1.}n_{.1}}{n}} = \frac{\frac{(nn_{11} - n_{1.}n_{.1})^2}{n}}{\frac{n_{1.}n_{.1}}{n}} = \frac{(nn_{11} - n_{1.}n_{.1})^2}{nn_{1.}n_{.1}} \quad (1)$$

$$X_{12}^2 = \frac{n_{12} - (\frac{n_{1.}n_{.2}}{n})}{\frac{n_{1.}n_{.2}}{n}} = \frac{\frac{(nn_{12} - n_{1.}n_{.2})^2}{n}}{\frac{n_{1.}n_{.2}}{n}} = \frac{(nn_{11} - n_{1.}n_{.2})^2}{nn_{1.}n_{.2}} \quad (2)$$

$$X_{13}^2 = \frac{n_{13} - (\frac{n_{1.}n_{.3}}{n})}{\frac{n_{1.}n_{.3}}{n}} = \frac{\frac{(nn_{13} - n_{1.}n_{.3})^2}{n}}{\frac{n_{1.}n_{.3}}{n}} = \frac{(nn_{13} - n_{1.}n_{.3})^2}{nn_{1.}n_{.3}} \quad (3)$$

$$X_{21}^2 = \frac{n_{21} - (\frac{n_{2.}n_{.1}}{n})}{\frac{n_{2.}n_{.1}}{n}} = \frac{\frac{(nn_{21} - n_{2.}n_{.1})^2}{n}}{\frac{n_{2.}n_{.1}}{n}} = \frac{(nn_{21} - n_{2.}n_{.1})^2}{nn_{2.}n_{.1}} \quad (4)$$

$$X_{22}^2 = \frac{n_{22} - (\frac{n_{2.}n_{.2}}{n})}{\frac{n_{2.}n_{.2}}{n}} = \frac{\frac{(nn_{22} - n_{2.}n_{.2})^2}{n}}{\frac{n_{2.}n_{.2}}{n}} = \frac{(nn_{22} - n_{2.}n_{.2})^2}{nn_{2.}n_{.2}} \quad (5)$$

$$X_{23}^2 = \frac{n_{23} - (\frac{n_{2.}n_{.3}}{n})}{\frac{n_{2.}n_{.3}}{n}} = \frac{\frac{(nn_{23} - n_{2.}n_{.3})^2}{n}}{\frac{n_{2.}n_{.3}}{n}} = \frac{(nn_{23} - n_{2.}n_{.3})^2}{nn_{2.}n_{.3}} \quad (6)$$

$$X_{31}^2 = \frac{n_{31} - (\frac{n_{3.}n_{.1}}{n})}{\frac{n_{3.}n_{.1}}{n}} = \frac{\frac{(nn_{31} - n_{3.}n_{.1})^2}{n}}{\frac{n_{3.}n_{.1}}{n}} = \frac{(nn_{31} - n_{3.}n_{.1})^2}{nn_{3.}n_{.1}} \quad (7)$$

$$X_{32}^2 = \frac{n_{32} - (\frac{n_{3.}n_{.2}}{n})}{\frac{n_{3.}n_{.2}}{n}} = \frac{\frac{(nn_{32} - n_{3.}n_{.2})^2}{n}}{\frac{n_{3.}n_{.2}}{n}} = \frac{(nn_{31} - n_{3.}n_{.2})^2}{nn_{3.}n_{.2}} \quad (8)$$

$$X_{33}^2 = \frac{n_{33} - (\frac{n_{3.}n_{.3}}{n})}{\frac{n_{3.}n_{.3}}{n}} = \frac{\frac{(nn_{33} - n_{3.}n_{.3})^2}{n}}{\frac{n_{3.}n_{.3}}{n}} = \frac{(nn_{33} - n_{3.}n_{.3})^2}{nn_{3.}n_{.3}} \quad (9)$$

Sum up the results (equation 1 to 9)

$$X^2 = \frac{(nn_{11} - n_{1.}n_{.1})^2}{nn_{1.}n_{.1}} + \frac{(nn_{11} - n_{1.}n_{.2})^2}{nn_{1.}n_{.2}} + \frac{(nn_{13} - n_{1.}n_{.3})^2}{nn_{1.}n_{.3}} + \frac{(nn_{21} - n_{2.}n_{.1})^2}{nn_{2.}n_{.1}} + \frac{(nn_{22} - n_{2.}n_{.2})^2}{nn_{2.}n_{.2}} + \frac{(nn_{23} - n_{2.}n_{.3})^2}{nn_{2.}n_{.3}} + \frac{(nn_{31} - n_{3.}n_{.1})^2}{nn_{3.}n_{.1}} + \frac{(nn_{31} - n_{3.}n_{.2})^2}{nn_{3.}n_{.2}} + \frac{(nn_{33} - n_{3.}n_{.3})^2}{nn_{3.}n_{.3}}$$

$$X^2 = \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{(nn_{ij} - n_{i.}n_{.j})^2}{nn_{i.}n_{.j}} \quad (10)$$

For r x c

$$X^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(nn_{ij} - n_{i.}n_{.j})^2}{nn_{i.}n_{.j}} \quad \text{i,j= 1,2,3.(r x c)} \quad (11)$$

Breaking equation (10) down to lower terms

$$\boldsymbol{X_{11}^2} = \frac{(nn_{11} - n_{1.}n_{.1})^2}{nn_{1.}n_{.1}} = \frac{(n_{11}(n_{11}+n_{12}+n_{13}+n_{21}+n_{22}+n_{23}+n_{31}+n_{32}+n_{33}) - (n_{11}+n_{12}+n_{13})(n_{11}+n_{21}+n_{31})^2}{nn_{1.}n_{.1}} =$$

$$\frac{(n_{11}^2 + n_{11}n_{12} + -n_{11}n_{13} + n_{11}n_{21} + n_{11}n_{22} + n_{11}n_{23} + n_{11}n_{31} + n_{11}n_{32} + n_{11}n_{33} - n_{11}^2 - n_{11}n_{21} - n_{11}n_{31} - n_{11}n_{12})^2 \; - n_{12}n_{21} - n_{12}n_{31} - n_{11}n_{13} - n_{13}n_{21} - n_{13}n_{31}}{nn_{1.}n_{.1}}$$

$$= \frac{(n_{11}n_{22} + n_{11}n_{32} + n_{11}n_{23} + n_{11}n_{33} - n_{12}n_{21} - n_{12}n_{31} - n_{13}n_{21} - n_{13}n_{31})^2}{nn_{1.}n_{.1}} \quad (12)$$

$$\boldsymbol{X_{12}^2} = \frac{(nn_{12} - n_{1.}n_{.2})^2}{nn_{1.}n_{.1}} = \frac{(n_{12}(n_{11}+n_{12}+n_{13}+n_{21}+n_{22}+n_{23}+n_{31}+n_{32}+n_{33}) - (n_{11}+n_{12}+n_{13})(n_{12}+n_{22}+n_{32})^2}{nn_{1.}n_{.2}} =$$

$$\frac{(-n_{11}n_{12} + -n_{12}^2 + -n_{12}n_{13} + n_{12}n_{21} + n_{12}n_{22} + n_{12}n_{23} + n_{12}n_{31} + n_{12}n_{32} + n_{12}n_{33} - n_{11}n_{12} - n_{11}n_{22} - n_{11}n_{32})^2 \; - n_{12}^2 - n_{12}n_{22} - n_{12}n_{32} - n_{12}n_{13} - n_{13}n_{22} - n_{13}n_{23}}{nn_{1.}n_{.2}}$$

$$= \frac{(n_{12}n_{21} + n_{12}n_{23} + n_{12}n_{31} + n_{12}n_{33} - n_{11}n_{22} - n_{11}n_{32} - n_{13}n_{22} - n_{13}n_{23})^2}{nn_{1.}n_{.2}} \quad (13)$$

$$\boldsymbol{X_{13}^2} = \frac{(nn_{13} - n_{1.}n_{.3})^2}{nn_{1.}n_{.3}} = \frac{(n_{13}(n_{11}+n_{12}+n_{13}+n_{21}+n_{22}+n_{23}+n_{31}+n_{32}+n_{33})-(n_{11}+n_{12}+n_{13})(n_{13}+n_{23}+n_{33})^2}{nn_{1.}n_{.3}} =$$

$$\frac{\begin{array}{c}(-n_{11}n_{13}+n_{12}n_{13}+n_{13}n_{21}+n_{13}n_{22}+n_{13}n_{23}+n_{13}n_{31}+-n_{13}^2+n_{13}n_{32}+n_{13}n_{33}-n_{11}n_{13}-n_{11}n_{23}-n_{11}n_{33})^2 \\ -n_{12}n_{13}-n_{12}n_{23}-n_{12}n_{33}-n_{13}^2-n_{13}n_{23}-n_{13}n_{33}\end{array}}{nn_{1.}n_{.3}}$$

$$=\frac{(n_{13}n_{21}+n_{13}n_{22}+n_{13}n_{31}+n_{13}n_{32}-n_{11}n_{23}-n_{11}n_{33}-n_{12}n_{23}-n_{12}n_{33})^2}{nn_{1.}n_{.3}} \quad (14)$$

$$\boldsymbol{X_{21}^2} = \frac{(nn_{21} - n_{2.}n_{.1})^2}{nn_{2.}n_{.1}} = \frac{(n_{21}(n_{11}+n_{12}+n_{13}+n_{21}+n_{22}+n_{23}+n_{31}+n_{32}+n_{33})-(n_{21}+n_{22}+n_{23})(n_{11}+n_{21}+n_{31})^2}{nn_{2.}n_{.1}} =$$

$$\frac{\begin{array}{c}(n_{11}n_{21}+n_{12}n_{21}+n_{13}n_{21}+n_{21}^2+n_{21}n_{22}+n_{21}n_{23}+n_{21}n_{31}+n_{21}n_{32}+n_{21}n_{33}-n_{11}n_{21}-n_{21}^2-n_{21}n_{31}-n_{11}n_{12})^2 \\ -n_{21}n_{22}-n_{22}n_{31}-n_{11}n_{23}-n_{23}n_{21}-n_{23}n_{31}\end{array}}{nn_{2.}n_{.1}}$$

$$=\frac{(n_{12}n_{21}+n_{13}n_{21}+n_{21}n_{32}+n_{21}n_{33}-n_{11}n_{22}-n_{22}n_{31}-n_{11}n_{23}-n_{23}n_{31})^2}{nn_{2.}n_{.1}} \quad (15)$$

$$\boldsymbol{X_{22}^2} = \frac{(nn_{22} - n_{2.}n_{.2})^2}{nn_{2.}n_{.2}} = \frac{(n_{22}(n_{11}+n_{12}+n_{13}+n_{21}+n_{22}+n_{23}+n_{31}+n_{32}+n_{33})-(n_{21}+n_{22}+n_{23})(n_{12}+n_{22}+n_{32})^2}{nn_{2.}n_{.2}} =$$

$$\frac{\begin{array}{c}(n_{11}n_{22}+n_{12}n_{22}+n_{13}n_{22}+n_{21}n_{22}+n_{22}^2+n_{22}n_{23}+n_{22}n_{31}+n_{22}n_{32}+n_{22}n_{33}-n_{12}n_{21}-n_{21}n_{22}-n_{21}n_{32})^2 \\ -n_{12}n_{22}-n_{22}^2-n_{22}n_{32}-n_{12}n_{23}-n_{23}n_{22}-n_{23}n_{31}\end{array}}{nn_{2.}n_{.1}}$$

$$=\frac{(n_{11}n_{22}+n_{13}n_{22}+n_{22}n_{31}+n_{22}n_{33}-n_{12}n_{21}-n_{21}n_{32}-n_{12}n_{23}-n_{23}n_{32})^2}{nn_{2.}n_{.2}} \quad (16)$$

$$\boldsymbol{X_{23}^2} = \frac{(nn_{23} - n_{2.}n_{.3})^2}{nn_{2.}n_{.3}} = \frac{(n_{23}(n_{11}+n_{12}+n_{13}+n_{21}+n_{22}+n_{23}+n_{31}+n_{32}+n_{33})-(n_{21}+n_{22}+n_{23})(n_{13}+n_{23}+n_{33})^2}{nn_{2.}n_{.3}} =$$

$$\frac{\begin{array}{c}(n_{11}n_{23}+n_{12}n_{23}+n_{13}n_{23}+n_{21}n_{23}+n_{22}n_{23}+n_{23}^2+n_{23}n_{31}+n_{23}n_{32}+n_{23}n_{33}-n_{13}n_{21}-n_{21}n_{23}-n_{21}n_{33})^2 \\ -n_{13}n_{22}-n_{22}n_{23}-n_{22}n_{23}-n_{13}n_{23}-n_{23}^2-n_{23}n_{33}\end{array}}{nn_{2.}n_{.1}}$$

$$=\frac{(n_{11}n_{23}+n_{12}n_{23}+n_{23}n_{31}+n_{23}n_{32}-n_{13}n_{21}-n_{21}n_{33}-n_{13}n_{22}-n_{22}n_{33})^2}{nn_{2.}n_{.3}} \quad (17)$$

$$\boldsymbol{X_{31}^2} = \frac{(nn_{31} - n_{3.}n_{.1})^2}{nn_{3.}n_{.1}} = \frac{(n_{31}(n_{11}+n_{12}+n_{13}+n_{21}+n_{22}+n_{23}+n_{31}+n_{32}+n_{33})-(n_{31}+n_{32}+n_{33})(n_{11}+n_{21}+n_{31}))^2}{nn_{3.}n_{.1}} =$$

$$\frac{\begin{array}{c}(-n_{11}n_{31}+n_{31}n_{12}+n_{31}n_{13}+n_{31}n_{21}+n_{31}n_{22}+n_{31}n_{23}+-n_{31}^2+n_{31}n_{32}+n_{31}n_{33}-n_{11}n_{31}-n_{21}n_{31}-n_{31}^2-n_{11}n_{32})^2 \\ -n_{32}n_{21}-n_{32}n_{31}-n_{11}n_{33}-n_{33}n_{21}-n_{33}n_{31}\end{array}}{nn_{3.}n_{.1}}$$

$$=\frac{(n_{12}n_{31}+n_{13}n_{31}+n_{22}n_{31}+n_{23}n_{31}-n_{11}n_{32}-n_{21}n_{32}-n_{11}n_{33}-n_{21}n_{33})^2}{nn_{3.}n_{.1}} \quad (18)$$

$$\boldsymbol{X_{32}^2} = \frac{(nn_{32} - n_{3.}n_{.2})^2}{nn_{3.}n_{.2}} = \frac{(n_{32}(n_{11}+n_{12}+n_{13}+n_{21}+n_{22}+n_{23}+n_{31}+n_{32}+n_{33})-(n_{31}+n_{32}+n_{33})(n_{12}+n_{22}+n_{32}))^2}{nn_{3.}n_{.2}} =$$

$$\frac{\begin{array}{c}(n_{11}n_{32}+n_{12}n_{32}+n_{13}n_{32}+n_{21}n_{32}+n_{22}n_{32}+n_{23}n_{32}+n_{31}n_{32}+-n_{32}^2+n_{32}n_{33}-n_{12}n_{31}-n_{31}n_{22}-n_{31}n_{32})^2 \\ -n_{32}^2-n_{12}n_{32}-n_{22}n_{32}-n_{12}n_{33}-n_{33}n_{22}-n_{32}n_{33}\end{array}}{nn_{3.}n_{.2}}$$

$$=\frac{(n_{11}n_{32}+n_{13}n_{32}+n_{21}n_{32}+n_{23}n_{32}-n_{12}n_{31}-n_{22}n_{31}-n_{12}n_{33}-n_{22}n_{33})^2}{nn_{3.}n_{.2}} \quad (19)$$

$$\boldsymbol{X_{33}^2} = \frac{(nn_{33} - n_{3.}n_{.3})^2}{nn_{3.}n_{.3}} = \frac{(n_{32}(n_{11}+n_{12}+n_{13}+n_{21}+n_{22}+n_{23}+n_{31}+n_{32}+n_{33})-(n_{31}+n_{32}+n_{33})(n_{13}+n_{23}+n_{33}))^2}{nn_{3.}n_{.3}} =$$

$$\frac{\begin{array}{c}(n_{11}n_{33}+n_{12}n_{33}+n_{13}n_{33}+n_{21}n_{33}+n_{22}n_{32}+n_{23}n_{33}+n_{31}n_{33}+n_{33}^2+n_{32}n_{33}-n_{13}n_{31}-n_{31}n_{23}-n_{31}n_{33})^2 \\ -n_{33}^2-n_{13}n_{32}-n_{23}n_{32}-n_{32}n_{33}-n_{33}n_{13}-n_{23}n_{33}\end{array}}{nn_{3.}n_{.3}}$$

$$=\frac{(n_{11}n_{33}+n_{12}n_{33}+n_{21}n_{33}+n_{22}n_{33}-n_{13}n_{31}-n_{23}n_{31}-n_{13}n_{32}-n_{23}n_{32})^2}{nn_{3.}n_{.3}} \quad (20)$$

Add equation 12 to 20

$$C^2 = \frac{(n_{11}(n_{22}+n_{32}+n_{23}+n_{33})-(n_{12}+n_{13})(n_{21}+n_{31}))^2}{nn_{1.}n_{.1}} + \frac{(n_{12}(n_{21}+n_{23}+n_{31}+n_{33})-(n_{11}+n_{13})(n_{22}+n_{32}))^2}{nn_{1.}n_{.2}} + \frac{(n_{13}(n_{22}+n_{21}+n_{31}+n_{32})-(n_{23}+n_{33})(n_{11}+n_{12}))^2}{nn_{1.}n_{.3}} + \frac{(n_{21}(n_{12}+n_{13}+n_{23}+n_{33})-(n_{22}+n_{23})(n_{11}+n_{31}))^2}{nn_{2.}n_{.1}} + \frac{(n_{22}(n_{11}+n_{31}+n_{23}+n_{33})-(n_{12}+n_{32})(n_{21}+n_{23}))^2}{nn_{2.}n_{.2}} + \frac{(n_{23}(n_{11}+n_{12}+n_{31}+n_{32})-(n_{21}+n_{22})(n_{13}+n_{33}))^2}{nn_{2.}n_{.3}} + \frac{(n_{31}(n_{12}+n_{13}+n_{22}+n_{23})-(n_{32}+n_{33})(n_{11}+n_{21}))^2}{nn_{3.}n_{.1}} + \frac{(n_{32}(n_{11}+n_{13}+n_{21}+n_{23})-(n_{12}+n_{22})(n_{31}+n_{33}))^2}{nn_{3.}n_{.2}} + \frac{(n_{33}(n_{11}+n_{12}+n_{22}+n_{23})-(n_{13}+n_{23})(n_{31}+n_{32}))^2}{nn_{3.}n_{.3}} \quad (21)$$

$$\mathrm{C}^2 = \sum_{i=1}^{3}\sum_{i=1}^{3}\frac{(C_{ij})^2}{n\,n_{i.}n_{.j}} \text{ where } C_{ij} = \det\begin{pmatrix} n_{ij} & n_{i.} \\ n_{.j} & n \end{pmatrix}$$

I, j= 1,2,3 (21)

For (r X c)

$$C^2 = \sum_{i=1}^{r}\sum_{i=1}^{c}\frac{(C_{ij})^2}{n\,n_{i.}n_{.j}} \text{ Where } C_{ij} = \det\begin{pmatrix} n_{ij} & n_{i.} \\ n_{.j} & n \end{pmatrix} \quad (22)$$

2.2 Co-factor method

C^2 or C square can also be derived directly from the categorical table.

For example from table 1 above

Table 1.1 showing the first cofactor elements

	A			Total
B	n_{11}	n_{12}	n_{13}	$n_{1.}$
	n_{21}	n_{22}	n_{23}	$n_{2.}$
	n_{31}	n_{32}	n_{33}	$n_{3.}$
Total	$n_{1.}$	$n_{2.}$	$n_{3.}$	n

$$C_{11}^2 \frac{(n_{11}(n_{22}+n_{32}+n_{23}+n_{33})-(n_{12}+n_{13})(n_{21}+n_{31}))^2}{nn_{1.}n_{.1}} \quad (23)$$

Table 1.2 showing the second cofactor elements

	A			Total
B	n_{11}	n_{12}	n_{13}	$n_{1.}$
	n_{21}	n_{22}	n_{23}	$n_{2.}$
	n_{31}	n_{32}	n_{33}	$n_{3.}$
Total	$n_{1.}$	$n_{2.}$	$n_{3.}$	n

$$C_{12}^2 = \frac{(n_{12}(n_{21}+n_{23}+n_{31}+n_{33})-(n_{11}+n_{13})(n_{22}+n_{32}))^2}{nn_{1.}n_{.2}} \quad (24)$$

Table 1.3 showing the third cofactor elements

	A			Total
B	n_{11}	n_{12}	n_{13}	$n_{1.}$
	n_{21}	n_{22}	n_{23}	$n_{2.}$
	n_{31}	n_{32}	n_{33}	$n_{3.}$
Total	$n_{1.}$	$n_{2.}$	$n_{3.}$	n

$$C_{13}^2 = \frac{(n_{13}(n_{22}+n_{21}+n_{31}+n_{32})-(n_{23}+n_{33})(n_{11}+n_{12}))^2}{nn_{1.}n_{.3}} \quad (25)$$

Table 1.4 showing the forth cofactor elements

	A			Total
B	n_{11}	n_{12}	n_{13}	$n_{1.}$
	n_{21}	n_{22}	n_{23}	$n_{2.}$
	n_{31}	n_{32}	n_{33}	$n_{3.}$
Total	$n_{1.}$	$n_{2.}$	$n_{3.}$	n

$$C_{21}^2 = \frac{(n_{21}(n_{12}+n_{13}+n_{23}+n_{33})-(n_{22}+n_{23})(n_{11}+n_{31}))^2}{nn_{2.}n_{.1}} \quad (26)$$

Table 1.5 showing the fifth cofactor elements

	A			Total
B	n_{11}	n_{12}	n_{13}	$n_{1.}$
	n_{21}	n_{22}	n_{23}	$n_{2.}$
	n_{31}	n_{32}	n_{33}	$n_{3.}$
Total	$n_{1.}$	$n_{2.}$	$n_{3.}$	n

$$C_{22}^2 = \frac{(n_{22}(n_{11}+n_{31}+n_{23}+n_{33})-(n_{12}+n_{32})(n_{21}+n_{23}))^2}{nn_{2.}n_{.2}} \quad (27)$$

Table 1.6 showing the sixth cofactor elements

	A			Total
B	n_{11}	n_{12}	n_{13}	$n_{1.}$
	n_{21}	n_{22}	n_{23}	$n_{2.}$
	n_{31}	n_{32}	n_{33}	$n_{3.}$
Total	$n_{1.}$	$n_{2.}$	$n_{3.}$	n

$$C_{23}^2 = \frac{(n_{23}(n_{11}+n_{12}+n_{31}+n_{32}) - (n_{21}+n_{22})(n_{13}+n_{33}))^2}{nn_{2.}n_{.3}} \quad (28)$$

Table 1.7 showing the seventh cofactor elements

	A			Total
B	n_{11}	n_{12}	n_{13}	$n_{1.}$
	n_{21}	n_{22}	n_{23}	$n_{2.}$
	n_{31}	n_{32}	n_{33}	$n_{3.}$
Total	$n_{1.}$	$n_{2.}$	$n_{3.}$	n

$$C_{31}^2 = \frac{(n_{31}(n_{12}+n_{13}+n_{22}+n_{23}) - (n_{32}+n_{33})(n_{11}+n_{21}))^2}{nn_{3.}n_{.1}} \quad (29)$$

Table 1.8 showing the eighth cofactor elements

	A			Total
B	n_{11}	n_{12}	n_{13}	$n_{1.}$
	n_{21}	n_{22}	n_{23}	$n_{2.}$
	n_{31}	n_{32}	n_{33}	$n_{3.}$
Total	$n_{1.}$	$n_{2.}$	$n_{3.}$	n

$$C_{32}^2 = \frac{(n_{32}(n_{11}+n_{13}+n_{21}+n_{23}) - (n_{12}+n_{22})(n_{31}+n_{33}))^2}{nn_{3.}n_{.2}} \quad (30)$$

Table 1.9 showing the ninth cofactor elements

	A			Total
B	n_{11}	n_{12}	n_{13}	$n_{1.}$
	n_{21}	n_{22}	n_{23}	$n_{2.}$
	n_{31}	n_{32}	n_{33}	$n_{3.}$
Total	$n_{1.}$	$n_{2.}$	$n_{3.}$	n

$$C_{33}^2 = \frac{(n_{33}(n_{11}+n_{12}+n_{22}+n_{23}) - (n_{13}+n_{23})(n_{31}+n_{32}))^2}{nn_{3.}n_{.3}} \quad (31)$$

Add equation 23 to equation 31

(from tables 1.1 to 1.9)

$$C^2 = \frac{(n_{11}(n_{22}+n_{32}+n_{23}+n_{33}) - (n_{12}+n_{13})(n_{21}+n_{31}))^2}{nn_{1.}n_{.1}} + \frac{(n_{12}(n_{21}+n_{23}+n_{31}+n_{33}) - (n_{11}+n_{13})(n_{22}+n_{32}))^2}{nn_{1.}n_{.2}} + \frac{(n_{13}(n_{22}+n_{21}+n_{31}+n_{32}) - (n_{23}+n_{33})(n_{11}+n_{12}))^2}{nn_{1.}n_{.3}} + \frac{(n_{21}(n_{12}+n_{13}+n_{23}+n_{33}) - (n_{22}+n_{23})(n_{11}+n_{31}))^2}{nn_{2.}n_{.1}} + \frac{(n_{22}(n_{11}+n_{31}+n_{23}+n_{33}) - (n_{12}+n_{32})(n_{21}+n_{23}))^2}{nn_{2.}n_{.2}} + \frac{(n_{23}(n_{11}+n_{12}+n_{31}+n_{32}) - (n_{21}+n_{22})(n_{13}+n_{33}))^2}{nn_{2.}n_{.3}} + \frac{(n_{31}(n_{12}+n_{13}+n_{22}+n_{23}) - (n_{32}+n_{33})(n_{11}+n_{21}))^2}{nn_{3.}n_{.1}} + \frac{(n_{32}(n_{11}+n_{13}+n_{21}+n_{23}) - (n_{12}+n_{22})(n_{31}+n_{33}))^2}{nn_{3.}n_{.2}} + \frac{(n_{33}(n_{11}+n_{12}+n_{22}+n_{23}) - (n_{13}+n_{23})(n_{31}+n_{32}))^2}{nn_{3.}n_{.3}}$$

$$= \sum_{i=1}^{3} \sum_{i=1}^{3} \frac{(C_{ij})^2}{n\, n_{i.} n_{.j}}$$

where $C_{ij} = \det \begin{pmatrix} n_{ij} & n_{i.} \\ n_{.j} & n \end{pmatrix}$ I, j= 1,2,3

3. R-PROGRAMMING CODE

```
Csquare=function(x)
{
 cs=c()
 res=c()
 r=nrow(x)
 c=ncol(x)
 ni.=rowSums(x)
 n.j=colSums(x)
 n=sum(ni.)
 for(i in 1:r)
   {
   cij=matrix(NA,2,2)
   cij[1,2]=ni.[i]
   for(j in 1:c)
   {
     cij[2,1]=n.j[j]
     cij[2,2]=n
     cij[1,1]=x[i,j]
     dt=det(cij)
     cs=c(cs,dt)
     rs=(dt^2)/(n*ni.[i]*n.j[j])
     res=c(res,rs)
     }
   }
   return(list("res"=res,"c=square"=sum(res)))
}
Csquare(x)
```

Data entering
X=matrix(c(x1,x2,x3,...xn), column dimension)
Press Csquare(x)

4. EMPIRICAL COMPARISON OF THE FORMER AND THE LATER

[Source: Gupta S. P. *Statistical method, volume 38, pp* 965-1001, 2009]

Table2: 1000 families were selected at random in a city to test the belief that high income families usually send their children to public schools and the low income families often sent their children to Government schools. The following results were obtained. [1]

B		SCHOOL		TOTAL
		PUBLIC	GOVT	
	LOW	370	430	800
	HIGH	130	70	200
Total		500	500	100

$$C^2 = \sum_i^2 \sum_j^2 \frac{\begin{vmatrix} n_{ij} & n_{i.} \\ n_{.j} & n \end{vmatrix}^2}{n n_{i.} n_{.j}} \qquad \text{i,j} = 1,2$$

$$C^2 = \frac{\begin{vmatrix} 370 & 800 \\ 500 & 1000 \end{vmatrix}^2}{500X800X1000} + \frac{\begin{vmatrix} 430 & 800 \\ 400 & 1000 \end{vmatrix}^2}{400X800X1000} + \frac{\begin{vmatrix} 130 & 200 \\ 500 & 1000 \end{vmatrix}^2}{500X200X1000} + \frac{\begin{vmatrix} 70 & 200 \\ 400 & 1000 \end{vmatrix}^2}{400X200X1000}$$

C^2= 2.25 + 37.8125 + 9.00 + 1.25=50.3125

Using former method

$$X^2 = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} (n_{ij} - e_{ij})^2}{e_{ij}}$$

$$e_{11} = \frac{500\ X\ 800}{1000} = 400 \quad e_{12} = \frac{400\ X\ 800}{1000} = 320$$
$$e_{21} = \frac{500\ X\ 200}{1000} = 100 \quad e_{22} = \frac{400\ X\ 200}{1000} = 80$$

$$X^2 = \frac{(370-400)^2}{400} + \frac{(430-320)^2}{320} + \frac{(130-100)^2}{100} + \frac{(70-80)^2}{80} =$$

= 2.25 + 37.8125 + 9.00 +1.25 =50.3125

[Source: Murray P. S. and Larry J.S. *theory and problems of statistics, fourth edition, pp* 294-315, 2008.]

Table3: The data in Table 3 were collected on how individuals prepared their taxes and their education level.

Tax Prep	Education Level			Total
	High School	Bachelor	Master	100
Computer	23	35	42	
Pen and paper	45	30	25	100
Total	68	65	67	200

$$C^2 = \sum_i^2 \sum_j^2 \frac{\begin{vmatrix} n_{ij} & n_{i.} \\ n_{.j} & n \end{vmatrix}^2}{n n_{i.} n_{.j}} \qquad \text{i=1,2 and j=1,2,3}$$

$$C^2 = \frac{\begin{vmatrix} 23 & 100 \\ 68 & 200 \end{vmatrix}^2}{100X68X200} + \frac{\begin{vmatrix} 35 & 100 \\ 65 & 200 \end{vmatrix}^2}{100X65X200} + \frac{\begin{vmatrix} 42 & 100 \\ 67 & 200 \end{vmatrix}^2}{100X67X200} + \frac{\begin{vmatrix} 45 & 100 \\ 68 & 200 \end{vmatrix}^2}{100X68X200} +$$

$$\frac{\begin{vmatrix} 30 & 100 \\ 65 & 200 \end{vmatrix}^2}{100X65X200} + \frac{\begin{vmatrix} 25 & 100 \\ 67 & 200 \end{vmatrix}^2}{100X67X200}$$

= 3.5588 + 0.1923 + 2.1567 +3.5588 + 0.1923 + 2.1567 = 11.8156

Using former method

$$X^2 = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} (n_{ij} - e_{ij})^2}{e_{ij}} \qquad \text{I= 1,2 and j=1,2,3}$$

$$e_{11} = \frac{100\ X\ 68}{200} = 34 \quad e_{12} = \frac{100\ X\ 65}{200} = 32.5$$
$$e_{13} = \frac{100\ X\ 67}{200} = 33.5 \quad e_{21} = \frac{100\ X\ 68}{200} = 34$$
$$e_{22} = \frac{100X\ 65}{200} = 32.5 \quad e_{23} = \frac{100\ X\ 67}{200} = 33.5$$

$$X^2 = \frac{(23-34)^2}{34} + \frac{(35-32.5)^2}{32.5} + \frac{(42-33.5)^2}{33.5} + \frac{(45-34)^2}{34} + \frac{(30-32.5)^2}{32.5} + \frac{(25-33.5)^2}{33.5}$$

= 3.559 +0.192 + 2.157 + 3.559 +0.192 + 2.157 = 11.816

Table4: Given (2X4) categorical data below

B	A				Total
	26	19	13	31	89
	17	12	23	28	80
Total	43	31	36	59	169

Using the suggested method

$$C^2 = \sum_i^2 \sum_j^2 \frac{\begin{vmatrix} n_{ij} & n_{i.} \\ n_{.j} & n \end{vmatrix}^2}{n n_{i.} n_{.j}} \qquad \text{i=1,2 and j=1,2,3,4}$$

$$C^2 = \frac{\begin{vmatrix} 26 & 89 \\ 43 & 169 \end{vmatrix}^2}{89X43X169} + \frac{\begin{vmatrix} 19 & 89 \\ 31 & 169 \end{vmatrix}^2}{31X89X169} + \frac{\begin{vmatrix} 13 & 89 \\ 36 & 169 \end{vmatrix}^2}{36X89X169} + \frac{\begin{vmatrix} 31 & 89 \\ 59 & 169 \end{vmatrix}^2}{31X59X169} +$$

$$\frac{\begin{vmatrix} 17 & 80 \\ 43 & 169 \end{vmatrix}^2}{43X80X169} + \frac{\begin{vmatrix} 12 & 80 \\ 31 & 169 \end{vmatrix}^2}{31X80X169} + \frac{\begin{vmatrix} 23 & 80 \\ 36 & 169 \end{vmatrix}^2}{36X80X169} + \frac{\begin{vmatrix} 28 & 80 \\ 59 & 169 \end{vmatrix}^2}{80X59X169}$$

=0.4971 + 0.4382 + 1.8782 + 0.00016 + 0.5530 + 0.4875 + 2.0834 + 0.00018 = 5.9322

Using the former method

$$X^2 = \frac{\sum_{i=1}^{r}\sum_{j=1}^{c}(n_{ij}-e_{ij})^2}{e_{ij}} \quad \text{I= 1,2 and j=1,2,3,4}$$

$$e_{11} = \frac{43 X 89}{169} = 22.6450 \quad e_{12} = \frac{31 X 89}{169} = 16.3254$$
$$e_{13} = \frac{36 X 89}{169} = 18.9586 \quad e_{14} = \frac{59 X 89}{169} = 31.0710$$
$$e_{21} = \frac{43X 80}{169} = 20.3550 \quad e_{22} = \frac{31 X 80}{169} = 14.6746$$

$$e_{23} = \frac{36 X 80}{169} = 17.0414 \quad e_{12} = \frac{59 X 80}{169} = 27.3290$$

$$X^2 = \frac{(26-22.6450)^2}{22.6450} + \frac{(19-16.3254)^2}{16.3254} + \frac{(13-18.9586)^2}{18.9586} + \frac{(31-31.0710)^2}{31.0710} + \frac{(17-20.3550)^2}{20.3550} + \frac{(12-14.6746)^2}{14.6746} + \frac{(23-17.0414)^2}{17.0414} + \frac{(28-27.3290)^2}{27.3290}$$

$= 0.4971 + 0.4875 + 1.8728 + 0.00016 + 0.5530 + 0.4875 + 2.0835 + 0.01647 = 5.9487$

Table 5: Summary of the results from Tables 2 to 4

Method	(2X2)	(2X3)	(2X4)
Former (chi-square)	50.3125	11.816	5.9487
Later (C-square)	50.3125	11.8156	5.9322

Calculating the Power

The power is calculated as follows:

1. Find *x*α such that $1-\chi^2(x_a/df)=\alpha$, where $\chi^2(x_a/df)$ is the area to the left of *x* under a Chi-square distribution with df degrees of freedom.
2. Power =1 − $\chi^2_{df,\lambda}$, where $\chi^2_{k,\lambda}$ is the left-tail area of the noncentral Chi-square distribution with *k* degrees of freedom and non-centrality parameter λ. Note that .$\lambda=Nw^2$ [8]

Table 6:The summary results of the power for chi-square and c-square values.

degree of freedom	C-square value	C-square beta	C-square power	chi-square value	chi-square beta	chi-square power
1	1.838235	0.175158034	0.824841966	1.839	0.175068275	0.824931725
1	21.70868	3.17352E-06	0.999996826	20.0589	7.50932E-06	0.999992491
2	13.52252	0.001157769	0.998842231	13.5225	0.001157781	0.998842219
2	13.05502	0.001462643	0.998537357	13.055	0.001462658	0.998537342
4	55.08785	3.11415E-11	1	55.0879	3.11407E-11	1
6	136.3834	5.80617E-27	1	136.3834	5.80617E-27	1

CONCLUSION

Apart from saving time by the later method in tables 2 to 4, it also reduces the risk of rejecting the null hypothesis due to too much approximation in the former method in table 6.
In the summary table above, it was discover that the statistic value from the later method is more accurate than the former method.

REFERENCES

[1] **Gupta S. P.** *Statistical method, volume 38, pp* 965-1001, 2009

[2] **Murray P. S. and Larry J.S.** *theory and problems of statistics, fourth edition, pp* 294-315, 2008.

[3] **Maura E. stokes and Charles S. Davis,** *Categorical data analysis using SAS system, second edition pp* 5-100, 2000.

[4] **Oyeyemi G.M, Adewara, A.A. Adebola F.B. and Salau S.I.**, *On the Estimation of Powerand Sample Size in Test of Independence. Asian Journal of Mathematics & Statistics* 3(3), 139-146. 2010

[5] **Lauritzen, Steffen L,** *Lecture on contingency tables (Pdf),(updated electronic version of the (University of Aalborg) 3rd* (1989), ed) 2002.

[6] **Yates, Daniel S; Moore David S; Starnes Daren S**. *The practice of statistics (2nd Ed) . New York Freeman*. ISBN 978-0-71674773-4. 2005.

[7] **Adam, Hayes**. *Math&Statistics (Investopedia)* 2020, http://www.investopedia.com>terms

[8] **John E.** *Power of chi-square test for 2x2 contingency tables with small expected frequencies. Psychology Bulletin vol* 87(1) 132-35 1980.