Equivalence of the *Linear-by-Linear* Chi-Square and the *N*-1 Chi-Square for 2×2 Tables

The equations shown below assume a 2×2 as follows:

	B = 0	*B* = 1	Total
A = 0	*a*	*b*	*m*
A = 1	*c*	*d*	*n*
Total	*r*	*s*	*N*

The SPSS CROSSTABS procedure computes a chi-square test that it labels *Linear-by-Linear Association*. The SPSS algorithms page for CROSSTABS describes it as the *Mantel-Haenszel Test of Linear Association*, and gives the formula shown in Equation 1, where *r* = Pearson's correlation.[1]

$$\chi^2_{MH} = (N-1)r^2 \tag{1}$$

For a 2×2 table, Pearson's chi-square can be computed using the formula shown in Equation 2.

$$\text{Pearson } \chi^2 = \frac{N(ad-bc)^2}{mnrs} \tag{2}$$

The N-1 chi-square is computed using that same formula, but with (*N*-1) in place of *N* in the numerator—see Equation 3.

$$(N-1) \text{ chi-square} = \frac{(N-1)(ad-bc)^2}{mnrs} \tag{3}$$

When Pearson's correlation is computed for two dichotomous variables, such as one has for a 2×2 table, it is often described as the *Phi* coefficient (r_ϕ). Before desktop computers and statistical software packages were readily available, r_ϕ and r_ϕ^2 were typically computed using the shortcuts shown in Equations 4 and 5.

$$r_\phi = \frac{ad-bc}{\sqrt{mnrs}} \tag{4}$$

$$r_\phi^2 = \frac{(ad-bc)^2}{mnrs} \tag{5}$$

Finally, multiplying the right side of Equation 5 by (*N*-1) yields Equation 3, which is the most common formula for the *N*-1 chi-square.

[1] It actually uses *W* in place of *N*, but the *W* stands for the total number of observations in the contingency table.

Putting it all together in a single equation, we get the following:

$$\chi^2_{MH} = (N-1)r^2 = \frac{(N-1)(ad-bc)^2}{mnrs} = \text{ the } N-1 \text{ chi-square} \quad (6)$$

Thus, for 2×2 tables, the *Linear-by-Linear Association* test computed by the SPSS CROSSTABS procedure is equivalent to the *N*-1 chi-square.

Acknowledgements

I thank Sacha Dubois for raising the question of whether the *Linear-by-Linear Association* chi-square was equivalent to the *N*-1 chi-square after noticing that they were the same for some data he was analyzing. His question prompted me to examine the formulae for the two measures. I also thank Ray Koopman for checking my work.