Brief Report

Applied Psychological Measurement
Volume 33 Number 6
September 2009 486-487

10.1177/0146621608320759
http://apm.sagepub.com
hosted at
http://online.sagepub.com

A Comment on "The *J* Index as a Measure of Nominal Scale Response Agreement"

Matthijs J. Warrens
Leiden University, The Netherlands

Various authors have proposed agreement indices for measuring nominal scale response agreement between two judges. Two situations may occur. Either the categories of the nominal scale are defined in advance and both raters use the same categories, or the categories are not defined in advance and the number of categories used by each rater is different. For the former case, the kappa statistic by Cohen (1960) is a popular measure. For the latter case, agreement measures have been proposed by Brennan and Light (1974), Hubert (1977), and Janson and Vegelius (1978, 1982).

Hubert's (1977) Γ is a monotonic function of the index by Brennan and Light (1974) and may be derived directly as a correlational index of agreement. Janson and Vegelius (1982) discussed some appealing properties of Hubert's Γ: It is a special case of Daniel-Kendall's generalized correlation coefficient, and it satisfies the requirement of a scalar product between normalized vectors in a Euclidean space. Janson and Vegelius (1982) also noted several less desirable characteristics of Hubert's Γ:

1. In general, Γ has a positive value when all frequencies are equal; the value zero would be preferable.
2. Γ has a negative value if both raters use only two categories and all frequencies are equal; the value zero would be preferable.
3. The minimum value of Γ is not zero; the value zero would be preferable.
4. If both raters use only two categories, Γ does not seem to be closely related to other association measures for dichotomous variables.

As an alternative to Hubert's Γ, Janson and Vegelius (1982) proposed a modified Γ (represented by Γ^*). Moreover, they investigated the above four characteristics for both Γ^* and the J index (Janson & Vegelius, 1978). The authors claimed that Γ exhibits all four, Γ^* two, and J none of the four undesirable characteristics. In this comment, it is shown that Γ^* only exhibits Characteristic 1, and not Characteristics 1 and 4 as is claimed by Janson and Vegelius (1982).

Let n_{11}, n_{12}, n_{21}, and n_{22} represent the four entries of a general 2×2 table, and let $n = n_{11} + n_{12} + n_{21} + n_{22}$. For two dichotomized variables, the three agreement indices are given by (Janson & Vegelius, 1982, Equations (8), (14), and (17)):

$$\Gamma = 1 - 4\frac{(n_{11} + n_{22})(n_{12} + n_{21})}{n(n+1)}, \tag{1}$$

$$\Gamma^* = 1 - 4\frac{(n_{11} + n_{22})(n_{12} + n_{21})}{n^2}, \tag{2}$$

and

$$J = \frac{[(n_{11} + n_{22}) - (n_{12} + n_{21})]^2}{n^2}. \quad (3)$$

Janson and Vegelius (1982) noted that Equations (1) and (2) do not seem to be closely related to other association measures for dichotomous variables. Instead, Equation (3) is equal to the square of the G index (Holley & Guilford, 1964). However, Equation (2) may be written as

$$\begin{aligned}
\Gamma^* &= \frac{n^2}{n^2} - 4\frac{(n_{11} + n_{22})(n_{12} + n_{21})}{n^2} \\
&= \frac{[(n_{11} + n_{22}) + (n_{12} + n_{21})]^2 - 4(n_{11} + n_{22})(n_{12} + n_{21})}{n^2} \\
&= \frac{(n_{11} + n_{22})^2 + (n_{12} + n_{21})^2 - 2(n_{11} + n_{22})(n_{12} + n_{21})}{n^2} \\
&= \frac{[(n_{11} + n_{22}) - (n_{12} + n_{21})]^2}{n^2} = J.
\end{aligned}$$

Thus, Γ^* and the J index are equivalent if both raters use only two categories. Γ^* proposed in Janson and Vegelius (1982) therefore exhibits only Characteristic 1.

For most types of data, multiple resemblance measures or association coefficients have been introduced. To choose the best or most appropriate coefficient, the various measures need to be better understood. Studying characteristics and special cases of resemblance measures (as is done in Janson & Vegelius, 1982), often provides us insight into the coefficients themselves. The fact that the J index and Γ^* are equivalent in the 2×2 case, suggests that the two measures have similar properties for the general nominal case, although their values are not the same in general. The behavior of J and Γ^* in the general nominal case is a topic for further investigation. For the moment, only J exhibits none of the four undesirable properties described above and may for that reason be preferred over Γ^*.

References

Brennan, R. L., & Light, R. J. (1974). Measuring agreement when two observers classify people into categories not defined in advance. *British Journal of Mathematical and Statistical Psychology*, *27*, 154-163.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37-46.

Holley, J. W., & Guilford, J. P. (1964). A note on the G index of agreement. *Educational and Psychological Measurement*, *24*, 749-753.

Hubert, L. J. (1977). Nominal scale response agreement as a generalized correlation. *British Journal of Mathematical and Statistical Psychology*, *30*, 98-103.

Janson, S., & Vegelius, J. (1978). On the applicability of truncated component analysis based on correlation coefficients for nominal scales. *Applied Psychological Measurement*, *2*, 135-145.

Janson, S., & Vegelius, J. (1982). The J index as a measure of nominal scale response agreement. *Applied Psychological Measurement*, *6*, 111-121.