ON FIXED POINTS OF THE CORRECTION FOR CHANCE FUNCTION FOR 2 X 2 ASSOCIATION COEFFICIENTS

Matthijs J. Warrens
Leiden University, Institute of Psychology, Unit Methodology and Statistics
P.O. Box 9555, 2300 RB Leiden, Email: warrens@fsw.leidenuniv.nl

ABSTRACT
This paper studies correction for chance for coefficients that are linear functions of the proportion of observed agreement. The fixed points of the correction for chance function are characterized. An equivalence relation on the set of linear functions is defined and it is shown that each linear function is mapped to the unique fixed point in its equivalence class.

keywords: *Chance-corrected coefficient; 2x2 coefficient; Intra-class kappa; Cohen's kappa.*

1. INTRODUCTION

Association coefficients are important entities in various domains of data analysis and classification. They are used to express the relationship between two variables in a number. In applications association coefficients are either used to summarize a particular research study, or they are used as input for multivariate data analysis techniques like, regression analysis, component analysis [7,8] or cluster analysis [1,15]. Four examples of association coefficients are, Pearson's product-moment correlation for measuring the linear dependence between two continuous variables, the Hubert-Arabie adjusted Rand index for comparing partitions of two different clustering algorithms [10,15,18], and the proportion of observed agreement and Cohen's kappa for assessing inter-rater agreement on a categorical scale [3,4,19,20,22].

In several data-analytic contexts it is desirable that the theoretical value of an association coefficient is zero if the two variables are statistically independent [12,23]. The Pearson correlation, the adjusted Rand index and Cohen's kappa each have zero value under independence, but the proportion of observed agreement does not. If a coefficient does not have zero value under statistical independence, it may be corrected for association due to chance [1,6,11,17]. After correction for chance a coefficient A has a form

$$c(A)=\frac{A-E(A)}{1-E(A)}, \tag{1}$$

where $E(A)$ is the value of coefficient A under chance. The 1 in the denominator of (1) is the maximum value of A. In this paper we only consider association coefficients with maximum value unity. The function c in (1) has been applied to, for example, association coefficients for metric scales [23,24], coefficients for interrater agreement [20,25], and coefficients for cluster validation [1,15].

Various authors have demonstrated that association coefficients may become equivalent after correction (1) [1,6,17,20,24]. These results deepen our understanding on how the various association coefficients that have been proposed in the literature are related, and provide new ways to interpret several important chance-corrected association coefficients. Here we are interested in c as a mathematical function. We study c in the context of association coefficients for 2×2 tables. This is not a severe limitation since many experimental and research studies can often be summarized by a 2×2 matrix or table [2,8,16,17]. This type of table is usually a cross-classification of two binary variables. An example from epidemiology is a reliability study in which two observers each rate the same sample of subjects on a the presence/absence of a trait [3,6]. An example from cluster analysis is a cluster validation study in which two partitions of the same set of points from two different clustering algorithms are compared [1,15,18].

The paper is organized as follows. In the next section we introduce the notation and definitions of association coefficients for 2×2 tables. In Section 3 we consider the correction for chance function. In Section 4 we show that the function c is idempotent, and we characterize the fixed points of c. Using an equivalence relation presented in Section 3 it is shown that each fixed point belongs to precisely one equivalence class and that the function c maps all elements of an equivalence class to the unique fixed point. Section 5 contains a conclusion.

2. ASSOCIATION COEFFICIENTS FOR 2 X 2 TABLES

In this section we introduce notation and definitions of association coefficients for 2×2 tables. A population 2×2

table is presented in Table 1. For notational convenience the table entries a, b, e and d are relative frequencies or proportions. The row totals p_1 and q_1 and column totals p_2 and q_2 are the marginal totals that result from summing the relative frequencies.

Table 1: Break-down of relative frequencies for two binary $(0,1)$ variables.

	Variable 2		
Variable 1	1	0	Totals
1	a	b	p_1
0	e	d	q_1
Totals	p_2	q_2	1

Association coefficients for 2×2 tables are here defined as functions from the set of all 2×2 matrices with non-negative real entries into the real numbers. We will use the set

$$M = \left\{ \begin{pmatrix} a & b \\ e & d \end{pmatrix} : a,b,e,d \geq 0, a+b+e+d = 1 \right\}$$

as the domain of the association coefficients. The requirement $a+b+e+d=1$ ensures that the entries are relative frequencies. A 2×2 association coefficient is then a function $A: M \to \mathbb{R}$, and the set of all such coefficients is denoted by $N = \{A: M \to \mathbb{R}\}$. Examples of elements of N are the odds ratio and the determinant, given by, respectively,

$$OR: \begin{pmatrix} a & b \\ e & d \end{pmatrix} \mapsto \frac{ad}{be}, \qquad and \qquad det: \begin{pmatrix} a & b \\ e & d \end{pmatrix} \mapsto ad - be.$$

Since a, b, e and d are proportions the determinant is equal to the covariance of two binary variables.

In this paper we are interested in 2×2 association coefficients that have a maximum value of 1. This excludes, for example, the odds ratio, since this coefficient has no upper bound. The determinant has maximum value 1/4, which is obtained when $a = d = 1/2$. Hence, the coefficient 4 ($ad - be$) is included. We also limit N to coefficients that are linear functions of the proportion of observed agreement $a+d$ given fixed marginal totals. The proportion of observed agreement, or the trace of the 2×2 table, is the proportion of 1s and 0s shared by the variables in the same positions. This coefficient is also known as the simple matching coefficient [14].

Let $\lambda = \lambda(p_1, p_2)$ and $\mu = \mu(p_1, p_2)$ be functions of the marginal totals p_1 and p_2. We will use the set

$$L = \{A: M \to \mathbb{R} : A = \lambda + \mu(a+d), A \leq 1\}$$

as the domain of the correction for chance formula in the next section. Due to the identity $a = d + p_1 - q_2$, linear in $a+d$ given the marginal totals is equivalent to linear in a and linear in d. The set of functions L has been studied in Warrens [16,17,21]. Albatineh et al [1] considered a similar family of cluster validation coefficients of the form $\alpha + \beta\sum_{i,j} m_{ij}^2$, where m_{ij} is the number of data points that are in cluster i according to the first clustering method and in cluster j according to the second clustering method. These authors studied what association coefficients coincide after correction for chance.

We consider three examples of elements of L.

Example 1. The phi coefficient

$$\phi = \frac{ad - be}{\sqrt{p_1 p_2 q_1 q_2}}$$

is the formula of Pearson's product-moment correlation coefficient for two binary variables. Pearson's correlation is widely used as a coefficient of linear dependence between two variables. We can write ϕ as $\lambda + \mu(a+d)$ where

$$\lambda = -\frac{p_1 p_2 + q_1 q_2}{2\sqrt{p_1 p_2 q_1 q_2}} \tag{2a}$$

$$\mu = \frac{1}{2\sqrt{p_1 p_2 q_1 q_2}}. \tag{2b}$$

Example 2. Let $r \in [0,1/2]$ be a weight and consider the function

$$S(r) = \frac{4(a+d) - \left(p_1^r p_2^{1-r} + p_1^{1-r} p_2^r\right)^2 - \left(q_1^r q_2^{1-r} + q_1^{1-r} q_2^r\right)^2}{4 - \left(p_1^r p_2^{1-r} + p_1^{1-r} p_2^r\right)^2 - \left(q_1^r q_2^{1-r} + q_1^{1-r} q_2^r\right)^2}, \tag{3}$$

where the quantity

$$\frac{p_1^r p_2^{1-r} + p_1^{1-r} p_2^r}{2} \tag{4}$$

is the Heinz mean of p_1 and p_2 [9]. Since $a+d \leq 1$ we have $S(r) \leq 1$. Furthermore, we can write $S(r)$ as $\lambda + \mu(a+d)$ where

$$\lambda = -\frac{\left(p_1^r p_2^{1-r} + p_1^{1-r} p_2^r\right)^2 + \left(q_1^r q_2^{1-r} + q_1^{1-r} q_2^r\right)^2}{4 - \left(p_1^r p_2^{1-r} + p_1^{1-r} p_2^r\right)^2 - \left(q_1^r q_2^{1-r} + q_1^{1-r} q_2^r\right)^2} \tag{5a}$$

$$\mu = \frac{4}{4 - \left(p_1^r p_2^{1-r} + p_1^{1-r} p_2^r\right)^2 - \left(q_1^r q_2^{1-r} + q_1^{1-r} q_2^r\right)^2}. \tag{5b}$$

Several coefficients from the literature are special cases of $S(r)$. Coefficient $S(0)$ is given by

$$\pi = \frac{4(a+d) - (p_1 + p_2)^2 - (q_1 + q_2)^2}{4 - (p_1 + p_2)^2 - (q_1 + q_2)^2}. \tag{6}$$

which is a coefficient proposed by Scott [13]. Coefficient (6) is also known as the intra-class kappa [3]. It is a standard tool for the analysis of agreement in a 2×2 reliability study. Coefficient $S(1/2)$ is given by

$$\kappa = \frac{a + d - p_1 p_2 - q_1 q_2}{1 - p_1 p_2 - q_1 q_2} = \frac{2(ad - be)}{p_1 q_2 + p_2 q_1}, \tag{7}$$

which is a coefficient proposed by Cohen [4]. Coefficient (7) is a popular association coefficient for summarizing the information in a cross-classification of two binary variables [16,17].

Example 3. Let $r \in [0,1]$ be a weight and consider the function

$$T(r) = \frac{ra + (1-r)d}{ra + \frac{1}{2}(b+e) + (1-r)d}. \tag{8}$$

This parameter family was first studied in Warrens [21]. Since

$$ra + (1-r)d = \left(r - \frac{1}{2}\right)(p_1 - q_2) + \frac{a+d}{2} \tag{9}$$

and

$$ra + \frac{b+e}{2} + (1-r)d = r(p_1 - q_2) + \frac{q_1 + q_2}{2},$$

we can write $T(r)$ as $\lambda + \mu(a+d)$ where

$$\lambda = \frac{\left(r - \frac{1}{2}\right)(p_1 - q_2)}{r(p_1 - q_2) + \frac{1}{2}(q_1 + q_2)} \tag{10a}$$

$$\mu = \frac{1}{2r(p_1 - q_2) + q_1 + q_2}. \tag{10b}$$

Several coefficients from the literature are special cases of $T(r)$. Coefficient $T(1/2)$ is the proportion of observed agreement or the simple matching coefficient [14]. It can be interpreted as the number of 1s and 0s shared by the variables in the same positions, divided by the total length of the variables. Coefficient $T(1)$ is the coefficient proposed in Dice [5], a widely used coefficient in ecological biology.

3. CORRECTION FOR CHANCE

Formula (1) presents the formula for a coefficient A after correction for chance. The value of $A \in L$ under chance, expectation $E(A)$, is a function of the marginal totals p_1 and p_2. More formally the function is given by

$$c : L \to L, A \mapsto \frac{A - E(A)}{1 - E(A)}, \tag{11}$$

where $E(A) < 1$ to avoid indeterminacy. Since E is a linear operator we have, for $A = \lambda + \mu(a+d)$, the identity $E(A) = \lambda + \mu E(a+d)$. Using this property Albatineh et al [1] showed that for $A \in L$ function c becomes

$$c : L \to L, \lambda + \mu(a+d) \mapsto \frac{a + d - E(a+d)}{\frac{1-\lambda}{\mu} - E(a+d)},$$

or simplified,

$$c(A) = \frac{a + d - E(a+d)}{\frac{1-\lambda}{\mu} - E(a+d)}. \tag{12}$$

The function c is a map from L to L if L is closed under c. Lemma 1 shows that this is the case.

Lemma 1.

L is closed under c.

Proof: Let $A \in L$ with $A = \lambda + \mu(a+d)$. The formula for $c(A)$ is presented in (12). Since $E(a+d)$ is a function of the marginal totals we can write $c(A)$ as $\lambda^* + \mu^*(a+d)$ where

$$\lambda^* = \frac{-E(a+d)}{\frac{1-\lambda}{\mu} - E(a+d)} \tag{13a}$$

$$\mu^* = \frac{1}{\frac{1-\lambda}{\mu} - E(a+d)}. \tag{13b}$$

Hence, $c(A) \in L$.

Formula (12) shows that elements of L coincide after correction for chance if they have the same ratio

$$\frac{1-\lambda}{\mu}. \tag{14}$$

This suggests the following definition. Two coefficients $A_1, A_2 \in L$ are said to be equivalent with respect to (12), denoted by $A_1 \sim A_2$, if they have the same ratio (14). It can be shown that $\sim$ is an equivalence relation on L. The equivalence relation divides the elements of L into equivalence classes, one class for each value of (14). We consider two examples of equivalence classes.

Example 4. For the phi coefficient in Example 1 ratio (14)

$$\frac{1-\lambda}{\mu} = p_1 p_2 + q_1 q_2 + 2\sqrt{p_1 p_2 q_1 q_2} = \left(\sqrt{p_1 p_2} + \sqrt{q_1 q_2}\right)^2. \tag{15}$$

Example 5. For parameter families $S(r)$ in (3) and $T(r)$ in (8) ratio (14)

$$\frac{1-\lambda}{\mu} = 1. \tag{16}$$

To obtain (16) we used (5) and (10) for $S(r)$ and $T(r)$ respectively. Hence, the special cases of $S(r)$ and $T(r)$ belong to the same equivalence class. Note that all special cases of $S(r)$ and all special cases of $T(r)$ coincide after correction (12), regardless of the values of r. This equivalence class is uncountably infinite.

Example 5 illustrates that the function c in (12) is many-to-one, and thus not injective. Since c is not injective it is not invertible.

Different definitions of $E(a+d)$ provide different versions of formula (12). We consider two examples of $E(a+d)$. Some other examples can be found in Warrens [17].

Example 6. We may assume that the data are a product of chance concerning two different frequency distributions with parameters p_1 and p_2 [4,11]. The expectation of an entry in Table 1 is defined by the product of the corresponding marginal totals. The expectation $E(a+d)$ is given by

$$E(a+d) = p_1 p_2 + q_1 q_2. \tag{17}$$

Expectation (17) is the value of $a+d$ under statistical independence. It can be obtained by considering all permutations of the observations of the first variable, while preserving the order of the observations of the second variable. If for each permutation the value of $a+d$ is calculated, then the arithmetic mean of these values is equal to $p_1 p_2 + q_1 q_2$.

Using (16) and (17) in (12) we obtain the coefficient κ in (7). Thus, all special cases of $T(r)$ in (8) are mapped to Cohen's κ if we use (17) [21]. Furthermore, all special cases of $S(r)$ in (3) are mapped to Cohen's κ if we use (17).

Example 7. We may assume that the frequency distribution with parameter p underlying the variables in Table 1 is the same for both variables. To estimate the parameter p we may use the Heinz mean of the marginal totals p_1 and p_2 in (4). This gives the expectation

$$E(a+d) = \left(\frac{p_1^r p_2^{1-r} + p_1^{1-r} p_2^r}{2}\right)^2 + \left(\frac{q_1^r q_2^{1-r} + q_1^{1-r} q_2^r}{2}\right)^2, \tag{18}$$

for $r \in [0, 1/2]$. Scott [13] and Krippendorff [11] consider the case $r = 0$, which corresponds to the arithmetic mean $(p_1 + p_2)/2$. In this case the expectation $E(a+d)$ is given by

$$E(a+d)=\left(\frac{p_1+p_2}{2}\right)^2+\left(\frac{q_1+q_2}{2}\right)^2. \tag{19}$$

Using (16) and (19) in (12) we obtain the coefficient π in (6). Thus, all special cases of $S(r)$ in (3) and $T(r)$ in (8) are mapped to Scott's π if we use (19). Furthermore, note that (18) allows us to formulate infinitely many versions of (12).

4. FIXED POINTS

In this section we consider the fixed points of c in (12). Lemma 2 shows that c is idempotent.

Lemma 2.

The function c is idempotent.

Proof: Using λ^* and μ^* in (13) we have

$$\frac{1-\lambda^*}{\mu^*}=\frac{1-\lambda}{\mu}.$$

Hence $c(c(A))=c(A)$.

Since c is a idempotent function it has at least one fixed point. We are interested what characterizes these fixed points. Let $\nu=\nu(p_1,p_2)$ be a function of the marginal totals p_1 and p_2, and consider the subset of L given by

$$F=\left\{A\in L: A=\frac{a+d-E(a+d)}{\nu}, \text{for some } \nu\neq 0\right\}.$$

Lemma 3 shows that F is the set of fixed points of c.

Lemma 3.

F is the set of fixed points of c.

Proof: ($\Rightarrow$) Let $A\in F$. Then

$$A=\frac{a+d-E(a+d)}{\nu} \quad \text{for some } \nu\neq 0.$$

Using

$$\lambda=\frac{-E(a+d)}{\nu} \quad \text{and} \quad \mu=\frac{1}{\nu}$$

in (12), we obtain

$$c(A)=\frac{a+d-E(a+d)}{\nu}.$$

Hence $c(A)=A$ and it follows that A is a fixed point.

($\Leftarrow$) Let $A\in L$ with $A=\lambda+\mu(a+d)$ be a fixed point. Then $A=c(A)$ or equivalently

$$\lambda+\mu(a+d)=\frac{a+d-E(a+d)}{\frac{1-\lambda}{\mu}-E(a+d)}.$$

Equating the $(a+d)$-parts and the `not'-$(a+d)$-parts on both sides of the equality, we obtain the identities

$$\lambda=\frac{-E(a+d)}{\frac{1-\lambda}{\mu}-E(a+d)} \quad \text{and} \quad \mu=\frac{1}{\frac{1-\lambda}{\mu}-E(a+d)}.$$

Setting $\nu = \dfrac{1-\lambda}{\mu} - E(a+d)$ we have

$$A = \lambda + \mu(a+d) = \frac{a+d-E(a+d)}{\nu}.$$

Hence, $A \in F$.

Lemma 3 shows that coefficients of the form

$$A = \frac{a+d-E(a+d)}{\nu} \quad \text{for some } \nu \neq 0, \tag{20}$$

are precisely the fixed points of c. For different definitions of $E(a+d)$ we have different versions of c and also different fixed points. Since ν in (20) can be any function, it follows that F is uncountably infinite, that is, c has infinitely many fixed points.
Lemma 4 shows that c maps each element of L not in F to an element of F.

Lemma 4.
Elements of L that are not fixed points are mapped to fixed points.
Proof: Suppose $A \in L$ is not a fixed point and let $c(A) = B$. Since c is idempotent we have $c(B) = c(c(A)) = c(A) = B$. Hence, B is a fixed point and A is mapped to a fixed point.

Lemma 4 shows that $c(L) = F$, that is, the image of c are the fixed points in L. Recall that ratio (14) divides the elements of L into equivalence classes. It follows from (12) that for a equivalence class with ratio $(1-\lambda)/\mu$ the fixed point is given by

$$A = \frac{a+d-E(a+d)}{\dfrac{1-\lambda}{\mu} - E(a+d)}.$$

Since the fixed point is unique, there is precisely one fixed point in each equivalent class. Thus, in each equivalence class c maps the elements to the unique fixed point.
It is not immediately clear that each equivalence class has infinitely many elements. Lemma 5 shows that for each fixed point we can construct an infinite family of coefficients that are in the same equivalence class. The function $Q(r)$ in (22) generalizes the function $T(r)$ in Example 3.

Lemma 5.
Let $A \in F$ with

$$A = \frac{a+d-E(a+d)}{\nu}. \tag{21}$$

Furthermore, let $r \in [0,1]$ and consider the parameter family

$$Q(r) = \frac{ra + (1-r)d}{(r - \frac{1}{2})(p_1 - q_2) + \frac{1}{2}\left(\nu + E(a+d)\right)}. \tag{22}$$

Then $c(Q(r)) = A$.
Proof: Using (9) we can write $Q(r)$ as $\lambda + \mu(a+d)$ where

$$\lambda = \frac{\left(r-\frac{1}{2}\right)(p_1-q_2)}{(r-\frac{1}{2})(p_1-q_2)+\frac{1}{2}\left(\nu+E(a+d)\right)} \tag{23a}$$

$$\mu = \frac{1}{2(r-\frac{1}{2})(p_1-q_2)+\nu+E(a+d)}. \tag{23b}$$

Using (23), ratio (14)

$$\frac{1-\lambda}{\mu} = \nu + E(a+d). \tag{24}$$

Using (24) in (12) we obtain (21).

5. CONCLUSION

In this paper we studied the correction for chance formula in the context of association coefficients for 2×2 tables. We focused on coefficients that are linear functions of the observed proportion of agreement $a+d$ given the marginal totals of the 2×2 table. For coefficients of the form $\lambda+\mu(a+d)$ there is a closed formula for correction for chance. The ratio $(1-\lambda)/\mu$ can be used to define an equivalence relation on the set of linear functions. It was shown that all coefficients in an equivalence class are mapped to the unique fixed point in the equivalence class. The image of the correction for chance function is the set of its fixed points. Furthermore, each equivalence class has infinitely many elements. In other words, in each equivalence class infinitely many 2×2 coefficients coincide after correction for chance.

It follows from the results in this paper that each 2×2 coefficient that has zero value under chance is a fixed point of some correction for chance function (Lemma 3). In the last section of the paper it was shown how to construct some 2×2 coefficients that are in the same equivalence class as the 2×2 coefficient that has zero value under chance. For example, for the phi coefficient

$$\phi = \frac{ad-be}{\sqrt{p_1p_2q_1q_2}} = \frac{a+d-p_1p_2-q_1q_2}{2\sqrt{p_1p_2q_1q_2}}$$

we have ratio (15). Let $r\in[0,1]$ and consider the coefficients

$$\frac{ra+(1-r)d}{(r-\frac{1}{2})(p_1-q_2)+\frac{1}{2}(\sqrt{p_1p_2}+\sqrt{q_1q_2})^2}$$

It follows from Lemma 5 that these coefficients become ϕ if $E(a+d) = p_1p_2+q_1q_2$.

6. ACKNOWLEDGMENT

This research was done while the author was funded by the Netherlands Organisation for Scientific Research, Veni project 451-11-026.

REFERENCES

[1]. A. N. Albatineh, M. Niewiadomska-Bugaj, and D. Mihalko. On similarity indices and correction for chance agreement. Journal of Classification, 23:301-313, 2006.

[2]. F. B. Baulieu. A classification of presence/absence based dissimilarity coefficients. Journal of Classification, 6:233-246, 1989.

[3]. D. A. Bloch and H. C. Kraemer. 2x2 kappa coefficients: Measures of agreement or association. Biometrics, 45:269-287, 1989.

[4]. J. Cohen. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20:37-46, 1960.

[5]. L. R. Dice. Measures of the amount of ecologic association between species. Ecology, 26:297-302, 1945.

[6]. J. L. Fleiss. Measuring agreement between two judges on the presence or absence of a trait. Biometrics, 31:651-659, 1975.

[7]. J. C. Gower. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika, 53:325-338, 1966.
[8]. J. C. Gower and P. Legendre. Metric and Euclidean properties of dissimilarity coefficients. Journal of Classification, 3:5-48, 1986.
[9]. E. Heinz. Beiträge zur Störungstheorie der Spektralzerlegung. Mathematische Annalen, 123:415-438, 1951.
[10]. L. J. Hubert and P. Arabie. Comparing partitions. Journal of Classification, 2:193-218, 1985.
[11]. K. Krippendorff. Association, agreement, and equity. Quality and Quantity, 21:109-123, 1987.
[12]. R. Popping. Overeenstemmingsmaten voor nominale data. Rijksuniversiteit Groningen, Groningen, 1983.
[13]. W. A. Scott. Reliability of content analysis: The case of nominal scale coding. Public Opinion Quarterly, 19:321-325, 1955.
[14]. R. R. Sokal and C. D. Michener. A statistical method for evaluating systematic relationships. University of Kansas Science Bulletin, 38:1409-1438, 1958.
[15]. D. Steinley. Properties of the Hubert-Arabie adjusted Rand index. Psychological Methods, 9:386-396, 2004.
[16]. M. J. Warrens. On association coefficients for 2x2 tables and properties that do not depend on the marginal distributions. Psychometrika, 73:777-789, 2008.
[17]. M. J. Warrens. On similarity coefficients for 2x2 tables and correction for chance. Psychometrika, 73:487-502, 2008.
[18]. M. J. Warrens. On the equivalence of Cohen's kappa and the Hubert-Arabie adjusted Rand index. Journal of Classification, 25:177-183, 2008.
[19]. M. J. Warrens. Cohen's kappa can always be increased and decreased by combining categories. Statistical Methodology, 7:673-677, 2010.
[20]. M. J. Warrens. Inequalities between kappa and kappa-like statistics for kxk tables. Psychometrika, 75:176-185, 2010.
[21]. M. J. Warrens. Chance-corrected measures for 2x2 tables that coincide with weighted kappa. British Journal of Mathematical and Statistical Psychology, 64:355-365, 2011.
[22]. M. J. Warrens. Conditional inequalities between Cohen's kappa and weighted kappas. Statistical Methodology, 10:14-22, 2013.
[23]. F. E. Zegers. A family of chance-corrected association coefficients for metric scales. Psychometrika, 51:559-562, 1986.
[24]. F. E. Zegers. A general family of association coefficients. Boomker, Groningen, 1986.
[25]. F. E. Zegers. Coefficients for interrater agreement. Applied Psychological Measurement, 15:321-333, 1991.