Kasetsart J. (Nat. Sci.) 43 : 389 - 397 (2009)

Some Statistical Aspects of Measuring Agreement Based on a Modified Kappa

Pornpis Yimprayoon[1]* and Sittipong Ruktamatakul[2]

ABSTRACT

The focus of this paper is the statistical inference of the problem of assessing agreement or disagreement between two raters who employ measurements on a two-level nominal scale. The purpose of this study was to derive the approximate asymptotic variance of the modified kappa statistic. Further, a comparison of the proposed estimate and an estimated large sample variance of Cohen's kappa is provided for all proportions expected to get a rating of 1 from each rater. When the value of the modified kappa is greater than or equal to 0.5 (or less than or equal to –0.5), the result of this study demonstrated that the sample estimate of the modified kappa is more efficient than the estimate of Cohen's kappa for each probability of being classified by both raters as category 1.

Key words: measuring agreement, Cohen's kappa, modified kappa, asymptotic mean, asymptotic variance

INTRODUCTION

Over the last decade, researchers have become increasingly aware of the problem of a methodology for measuring agreement on assessing the acceptability of a new or generic process which can arise throughout many scientific and non-scientific fields. Measuring agreement has been used very often to designate the level of agreement between different data-generating sources referred to as raters. A rater could be a clinician, a nurse, a psychologist, a radiologist, a chemist, a statistician, a pharmacist, a laboratory apparatus, an instrument, a rating system, a diagnosis, a treatment, a method, a process, a technique or a formula. There are numerous examples that illustrate these situations. Firstly, in education and social science measurement, the comparison of a newly developed measurement method with an established one is often used to see whether there is sufficient agreement for the new to replace the old. This makes sure that the new method of measurement is cheap, quick, correct and optimal. Secondly, in clinical and medical diagnosis problems, a team of physicians is used in order to diagnose and select the appropriate treatment for a comatose patient. Thirdly, in criminal trials, a group of jurors is used and sentencing depends on complete agreement among the jurors. Fourthly, hotels receive five stars only after several visitors agree on the service. Finally, the medals and rankings in sport games are based on the ratings of several judges.

One of the most popular indices of

[1] Department of Mathematics, Faculty of Liberal Arts and Science, Kasetsart University, Kamphaeng Saen, Nakhon Pathom 73140, Thailand.

[2] Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Rangsit Center, Prathum Thani 12121, Thailand.

* Corresponding author, e-mail: faasppy@ku.ac.th

Received date : 09/12/08 Accepted date : 04/02/09

agreement was originally presented by Cohen (1960), namely Cohen's kappa statistic (κ_C), as a reliability index for measuring agreement between two raters employing nominal scales. Later, in 1968, Cohen generalized the original kappa statistic to a weighted kappa that provided for the incorporation of a ratio-scaled degree of disagreement (or agreement) to each of the cells of the $k \times k$ table of joint nominal scale assignments, such that disagreements of varying gravity (or agreements of varying degree) were weighted accordingly.

Formulas for estimators of standard errors for kappa can be found, for instance, in Hildebrand *et al.* (1977) and in Liebetrau (1983). When the sample size is sufficiently large, Everitt (1968) and Fleiss *et al.* (1969) gave valid formulas for the approximate, large-sample mean and variance of the two statistics, kappa and weighted kappa.

Moreover, in the work of Landis and Koch (1977), it was found that weighted kappa was appropriate for measuring agreement when the categories of response were ordinal. Landis and Koch (1977) also proposed an approach by expressing the quantities which reflected the extent to which the raters agreed among themselves as functions of observed proportions obtained from underlying, multidimensional contingency tables. Davies and Fleiss (1982) proposed a generalization for multiple observers by the average of pairwise agreement. Some limitations of the kappa index are known, for example, that its value depends on the balance and symmetry of the marginal totals of the table (Feinstein and Cicchetti,1990; Guggenmoos-Holzmann, 1993) and some alternative methods of evaluating agreement among observers have been proposed (Donner and Donald, 1988; Cicchetti and Feinstein, 1990; Graham and Jackson, 1993). Abraira and Pérez de Vargas (1999) generalized the proposals of Schouten (1986) and Gross (1986) for multiple observers and incomplete design, in order to encompass ordinal variables with the inclusion of weights to enable pondering the severity of disagreement among different categories.

Several authors have proposed guidelines for the interpretation of the kappa statistic (Fleiss, 1981; Lantz and Nebenzahl, 1996). Even a matter as simple as the range of kappa is not clear from the literature but many discussions of kappa state that it ranges from –1 to 1, with 0 indicating no agreement beyond that expected by chance and 1 indicating perfect agreement. A comprehensive review paper is provided by Banerjee *et al.* (1999).

The use of kappa for qualitative or dichotomous judgments, such as presence or absence of disease, has been described by Fleiss and Chilton (1983). The kappa index is still a very frequently used statistic in clinical epidemiological literature (e.g. Jelles *et al.*, 1995; Pérez *et al.*, 1997). In addition, many other applications of the kappa statistic in a variety of different contexts can be found in the recent works of Guimarães *et al.* (2008), Jittavisutthikul *et al.* (2008) and Prabhasavat and Homgade (2008).

MATERIALS AND METHODS

Brief description of Cohen's kappa statistic

Consider a reliability research where two raters, referred to as rater A and rater B, are required to classify n subjects into one of two possible response categories. The two response categories, labeled as 1 and 2, are assumed to be disjoint. Denote π_{ij} as the chance that rater A classifies a subject into category i, while rater B classifies the same subject into category j, i,j=1,2.

Let $\pi_{1\cdot} = \sum_{j=1}^{2} \pi_{1j}$ and $\pi_{2\cdot} = \sum_{j=1}^{2} \pi_{2j}$ be the probability of being classified by rater A to categories 1 and 2, respectively. The probabilities $\pi_{\cdot 1} = \sum_{i=1}^{2} \pi_{i1}$ and $\pi_{\cdot 2} = \sum_{i=1}^{2} \pi_{i2}$ are also defined in the same manner.

Under these conditions, Cohen's kappa statistic for measuring agreement between the two raters is defined as

$$\kappa_C = \frac{\theta_o - \theta_e}{1 - \theta_e} \quad (1)$$

where

$$\theta_o = \pi_{11} + \pi_{22},\ \theta_e = \pi_{1\cdot}\pi_{\cdot 1} + \pi_{2\cdot}\pi_{\cdot 2}. \quad (2)$$

In applications, if there are n subjects and n_{ij} represents the number of subjects classified in category i by rater A and in category j by rater B, the sample estimate of κ_C is given by

$$\hat{\kappa}_C = \frac{\hat{\theta}_o - \hat{\theta}_e}{1 - \hat{\theta}_e} \quad (3)$$

where

$$\hat{\pi}_{ij} = \frac{n_{ij}}{n}, \hat{\pi}_{i\cdot} = \frac{n_{i\cdot}}{n}, \hat{\pi}_{\cdot j} = \frac{n_{\cdot j}}{n},$$

$$\hat{\theta}_o = \frac{n_{11} + n_{22}}{n}, \hat{\theta}_e = \frac{n_{1\cdot}n_{\cdot 1} + n_{2\cdot}n_{\cdot 2}}{n^2} \quad (4)$$

Limitations of Cohen's kappa statistic

Sinha *et al*. (2006) critically examined some features of κ_C. The properties of κ_C are:

(i) $\kappa_C = 1$ if and only if $\theta_o=1$. This means that there are no controversial judgments by the two raters i.e., the disagreement cells [(1, 2) and (2,1)] have zero probability each.

(ii) $\kappa_C = 0$ if and only if $\theta_o = \theta_e$. Technically, this holds if and only if

$$(\pi_{11} - \pi_{1\cdot}\pi_{\cdot 1}) + (\pi_{22} - \pi_{2\cdot}\pi_{\cdot 2}) \quad (5)$$

which, in its turn, implies that

$$\pi_{ij} = \pi_{i\cdot}\pi_{\cdot j};\ i,j = 1,2. \quad (6)$$

(iii) $\kappa_C = -1$ if and only if $\pi_{11} + \pi_{22} = 0$, $\pi_{12} = \pi_{21} = 0.5$. Technically, this means that both the agreement cells have zero probability each, while the two disagreement cells are equally likely.

Sinha *et al*. (2006) pointed out some undesirable features of κ_C and also said that the case of "$\kappa_C = -1$" seemed to restrict behavior on the part of the raters. When $\pi_{11} = \pi_{22} = 0$, there is already an indication of total disagreement between the two raters. Therefore, in such situations, irrespective of the values assumed by π_{12} and π_{21} ($0<\pi_{12}, \pi_{21}<1, \pi_{12}+\pi_{21}=1$) it is desired that the kappa coefficient assumes the value -1. With this in mind, they set $\pi_{12} = \alpha$ and $\pi_{21} = 1-\alpha$, $0<\alpha<1$ and analyzed the situation with the purpose of modifying the definition of κ_C to deal with the full strength of disagreement between the two raters, while the ratings are given independently in a two-point nominal scale.

Their modification was aimed at the value $\kappa_C = -1$ They modified κ_C as

$$\kappa_M = \frac{\theta_o - \theta_e}{A - \theta_e} \quad (7)$$

and suggested a value of A to take care of the situations:

$$\pi_{11} = \pi_{22} = 0,\ \pi_{12} = \alpha,$$
$$\pi_{21} = 1-\alpha,\ 0 < \alpha < 1 \quad (8)$$

along with $\kappa_M = -1$. Under (8), κ_M reduces to

$$\kappa_M = \frac{-2\alpha(1-\alpha)}{A - 2\alpha(1-\alpha)} \quad (9)$$

and $\kappa_M = -1$ yields

$$A = 4\alpha(1-\alpha). \quad (10)$$

Then, replacing α by $\frac{\pi_{1\cdot} + \pi_{\cdot 2}}{2}$ in (10) produces

$$A = 4 \cdot \frac{\pi_{1\cdot} + \pi_{\cdot 2}}{2} \cdot \frac{\pi_{\cdot 1} + \pi_{2\cdot}}{2}$$
$$(\pi_{1\cdot} + \pi_{\cdot 2})(\pi_{\cdot 1} + \pi_{2\cdot}). \quad (11)$$

Next, substituting (11) in (7) produces

$$\kappa_M = \frac{\theta_o - \theta_e}{(\pi_{1\cdot} + \pi_{\cdot 2})(\pi_{\cdot 1} + \pi_{2\cdot}) - (\pi_{1\cdot}\pi_{\cdot 1} + \pi_{\cdot 2}\pi_{2\cdot})}. \quad (12)$$

Hence, the modified kappa statistic κ_M is defined as

$$\kappa_M = \frac{\theta_o - \theta_e}{\pi_{1.}\pi_{2.} + \pi_{.1}\pi_{.2}}. \qquad (13)$$

This modification is based on the analysis of situations leading to total disagreement between the two raters and all of the three essential features of the kappa statistic are retained by κ_M.

The current study applies and extends the work of Sinha *et al.* (2006) by deriving the large sample variance of the modified kappa statistic $V(\hat{\kappa}_M)$. It is necessary to propose $V(\hat{\kappa}_M)$ in order to see whether the modified Cohen's kappa statistic can be used to replace Cohen's kappa statistic by comparing the estimate of the asymptotic variance of $\hat{\kappa}_M$ against the variance estimate of $\hat{\kappa}_C$.

RESULTS AND DISCUSSION

From (13), the estimate of κ_M can be obtained by

$$\hat{\kappa}_M = \frac{n(n_{11}+n_{22})-(n_{11}+n_{12})(n_{11}+n_{21})-(n_{21}+n_{22})(n_{12}+n_{22})}{(n_{11}+n_{12})(n_{21}+n_{22})+(n_{11}+n_{21})(n_{12}+n_{22})} \qquad (14)$$

When the sample is sufficiently large (n>30), large sample theory can be used to evaluate the expected value and variance of $\hat{\kappa}_M (E(\hat{\kappa}_M))$. It then can be shown that the asymptotic mean of $\hat{\kappa}_M$ is κ_M, that is $E(\hat{\kappa}_M) = \kappa_M$, and the approximate asymptotic variance expression of $\hat{\kappa}_M$ is given by

$$V\left(g(n_{11},n_{12},n_{21})\right) = \begin{pmatrix} \frac{\partial g}{\partial n_{11}} & \frac{\partial g}{\partial n_{12}} & \frac{\partial g}{\partial n_{21}} \end{pmatrix} \cdot$$

$$n\begin{pmatrix} \pi_{11}(1-\pi_{11}) & -\pi_{11}\pi_{12} & -\pi_{11}\pi_{21} \\ -\pi_{11}\pi_{12} & \pi_{12}(1-\pi_{12}) & -\pi_{12}\pi_{21} \\ -\pi_{11}\pi_{21} & -\pi_{12}\pi_{21} & \pi_{21}(1-\pi_{21}) \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial g}{\partial n_{11}} \\ \frac{\partial g}{\partial n_{12}} \\ \frac{\partial g}{\partial n_{21}} \end{pmatrix} \qquad (15)$$

or equivalently,

$$V\left(g(n_{11},n_{12},n_{21})\right) \approx \Bigg[V(n_{11})\left(\frac{\partial g}{\partial n_{11}}\right)^2 + V(n_{12})\left(\frac{\partial g}{\partial n_{12}}\right)^2 + V(n_{21})\left(\frac{\partial g}{\partial n_{21}}\right)^2 + 2Cov(n_{11},n_{12})\left(\frac{\partial g}{\partial n_{11}}\right)\left(\frac{\partial g}{\partial n_{12}}\right) + 2Cov(n_{11},n_{21})\left(\frac{\partial g}{\partial n_{11}}\right)\left(\frac{\partial g}{\partial n_{21}}\right) + 2Cov(n_{12},n_{21})\left(\frac{\partial g}{\partial n_{12}}\right)\left(\frac{\partial g}{\partial n_{21}}\right) \Bigg]_{\substack{n_{11}\to n\pi_{11} \\ n_{12}\to n\pi_{12} \\ n_{21}\to n\pi_{21}}} \qquad (16)$$

where

$$g(n_{11},n_{12},n_{21}) = \frac{2\left(n_{11}^2 + n_{12}n_{21} + n_{11}(-n+n_{12}+n_{21})\right)}{2n_{11}^2 - nn_{12} + n_{12}^2 + n_{21}(-n+n_{21}) + 2n_{11}(-n+n_{12}+n_{21})} \qquad (17)$$

That is,

$$\begin{aligned} V(\hat{\kappa}_M) = {} & 4[2\pi_{11}^3(-1+\pi_{12}+\pi_{21})^2\{4\pi_{12}^2 + \\ & \pi_{21}(-1+4\pi_{21}) - \pi_{12}(1+8\pi_{21})\} \\ & +\pi_{12}\pi_{21}\{\pi_{12}^5 + \pi_{12}^4(-2+\pi_{21}) - \\ & 2\pi_{12}^2\pi_{21}^2(-1+\pi_{21}) + \pi_{12}\pi_{21}^3(-1+\pi_{21}) \\ & +\pi_{21}^3(-1+\pi_{21})^2 - \pi_{12}^3(-1+\pi_{21}+2\pi_{21}^2)\} \\ & +\pi_{11}^4\{4\pi_{12}^3 - \pi_{12}^2(5+4\pi_{21}) + \pi_{12}(1+6\pi_{21}-4\pi_{21}^2) \\ & +\pi_{21}(1-5\pi_{21}+4\pi_{21}^2)\} + \pi_{11}\{\pi_{12}^6 + \pi_{21}^2(-1+\pi_{21})^4 \\ & +\pi_{12}^5(-4+6\pi_{21}) - \pi_{12}^4(-6+10\pi_{21}+\pi_{21}^2) \\ & +2\pi_{12}\pi_{21}(-1+\pi_{21})^2(1+\pi_{21}+3\pi_{21}^2) \\ & +\pi_{12}^3(-4+4\pi_{21}+6\pi_{21}^5-12\pi_{21}^3) \\ & -\pi_{12}^2(-1+2\pi_{21}+4\pi_{21}^2-6\pi_{21}^3+\pi_{21}^4)\} \\ & +\pi_{11}^2\{5\pi_{12}^5 + \pi_{21}(-1+\pi_{21})^3(-1+5\pi_{21}) + \pi_{12}^4(-16+9\pi_{21}) \\ & -2\pi_{12}^3(-9+3\pi_{21}+7\pi_{21}^3) - 2\pi_{12}^2(4+2\pi_{21}-10\pi_{21}^2+7\pi_{21}^3) \\ & +\pi_{12}(1-4\pi_{21}^2-6\pi_{21}^3+9\pi_{21}^4)\}] \big/ \\ & [n\{2\pi_{11}^2 - \pi_{12} + \pi_{12}^2 + \pi_{21}(-1+\pi_{21}) \\ & +2\pi_{11}(-1+\pi_{12}+\pi_{21})\}^4]. \end{aligned} \qquad (18)$$

Finally, upon simplification, the asymptotic variance of $\hat{\kappa}_M$ can be written in the form

$$V(\hat{\kappa}_M) = \frac{Q_M}{n} \quad (19)$$

where

$$Q_M = \frac{1}{(\pi_{1.}\pi_{2.} + \pi_{.1}\pi_{.2})^2}$$

$$\sum_{i \neq j}^{2}\Big[\pi_{ii}\left\{(\pi_{jj} - \pi_{ii}) - \kappa_M(1 + \pi_{jj} - \pi_{ii})\right\}$$

$$\left\{(1+\kappa_M)\left(\theta_o(1-\theta_o) + 4\pi_{ij}\pi_{ji}\right)\right.$$

$$+(\pi_{jj} - \pi_{ii})(1 + \pi_{jj} - \pi_{ii})$$

$$\left.-\kappa_M(1 - \pi_{ii} + 2\pi_{jj} + \pi_{ii}^2 - \pi_{jj}^2)\right\}$$

$$-\pi_{ij}(1+\kappa_M)(\theta_o + 2\pi_{ij})$$

$$\{2\pi_{ij}(2\pi_{ij} - 1) + \theta_o(1-\theta_o)\}\Big] \quad (20)$$

A comparison of $\hat{\kappa}_M$ and $\hat{\kappa}_C$ is of interest and can be carried out by comparing the asymptotic variance of $\hat{\kappa}_M$ with the asymptotic variance of $\hat{\kappa}_C$ Base on Fleiss *et al.* (1969), the approximate asymptotic expression for the variance of $\hat{\kappa}_C$ can be given by

$$V(\hat{\kappa}_C) = \frac{Q_C}{n} \quad (21)$$

where, in general terms, Q_C is defined as

$$Q_C = \frac{1}{(1-\theta_e)^2}\left[\sum_{i=1}^{2}\pi_{ii}\left\{1-(\pi_{i.}+\pi_{.i})(1-\kappa_C)\right\}^2\right.$$

$$\left.+(1-\kappa_C)^2\sum_{i\neq j}^{2}\pi_{ij}(\pi_{i.}-\pi_{.j})^2 - \left\{\kappa_C - \theta_e(1-\kappa_C)\right\}^2\right] \quad (22)$$

It has further been noted by Cantor (1996) that all of the values $\pi_{2.}$, $\pi_{.2}$, θ_e, θ_o, π_{22}, π_{11}, π_{12}, and π_{21} can be determined by $\pi_{1.}$, $\pi_{.1}$ and κ_C, namely

$$\pi_{2.} = 1-\pi_{1.}, \pi_{.2} = 1-\pi_{.1}, \theta_e = \pi_{1.}\pi_{.1} + \pi_{2.}\pi_{.2},$$
$$\theta_o = \kappa_C(1-\theta_e) + \theta_e, \pi_{22} = (\theta_o - \pi_{1.} + \pi_{.2})/2,$$
$$\pi_{11} = \theta_o - \pi_{22}, \pi_{12} = \pi_{1.} - \pi_{11},$$
$$\text{and } \pi_{21} = \pi_{.1} - \pi_{11}. \quad (23)$$

Since $\theta_o - \theta_e = \kappa_M(\pi_{1.}\pi_{2.} + \pi_{.1}\pi_{.2})$, it turns out that an analogous result also holds in terms of $\pi_{1.}$, $\pi_{.1}$ and κ_M. Therefore, it is now possible to use the comparison of Q_M and Q_C instead of the comparison of the asymptotic variance of $\hat{\kappa}_M$ and the estimated large sample variance of $\hat{\kappa}_C$. For values of $\pi_{.1}$ and $\pi_{1.}$, Tables 1 and 2 display the values of Q_C for κ_C and the values of Q_M for κ_M, respectively. Since the upper bounds of κ_C and κ_M are less than 1 for $\pi_{1.} \neq \pi_{.1}$, Tables 1 and 2 thus have no values of Q_C for κ_C and Q_M for κ_M that are not permissible.

It follows from Tables 1 and 2 that, based on the same value of κ_M and κ_C when $\kappa_M \leq$-0.5 (or $\kappa_C \leq$-0.5) and $\kappa_M \geq 0.5$ (or $\kappa_C \geq 0.5$), Q_M is less than Q_C in all values of $\pi_{1.}$ and $\pi_{.1}$. This implies that the proposed estimator $V(\hat{\kappa}_M)$ is more efficient than $V(\hat{\kappa}_M)$ when $(\kappa_M)^2 \geq 0.5$ (or $(\kappa_C)^2 \geq 0.5$).

CONCLUSION

This paper discussed the problem of measuring agreement or disagreement between two raters where the ratings are given separately in a two-point nominal scale. The discussion focused on the modified Cohen's kappa statistic κ_M which deals with the full strength of disagreement between the two raters. In addition, a method was proposed to determine the estimated asymptotic variance of the modified Cohen's kappa statistic. Based on the comparison of the proposed estimate $V(\hat{\kappa}_M)$ with the estimated large sample variance of Cohen's kappa $V(\hat{\kappa}_C)$, it was concluded that, for $(\kappa_M)^2 \geq 0.5$ (or $(\kappa_C)^2 \geq 0.5$), the preference is mostly for κ_M in all values of $\pi_{1.}$ and $\pi_{.1}$.

ACKNOWLEDGEMENTS

The authors offer their sincere thanks to Assoc. Professor Dr Montip Tiensuwan for suggesting this problem. A special gratitude is

Table 1 Values of Q_C for κ_C.

$\pi_{1.}$	$\pi_{.1}$	κ_C																		
		-0.9	**-0.8**	**-0.7**	**-0.6**	**-0.5**	**-0.4**	**-0.3**	**-0.2**	**-0.1**	**0.0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
0.1	**0.1**									1.690	1.000	1.598	1.984	2.179	2.205	2.083	1.835	1.481	1.043	0.542
0.1	**0.2**									0.417	0.852	1.159	1.350	1.434	1.425	1.331	1.166			
0.1	**0.3**									0.430	0.654	0.808	0.899	0.931	0.911					
0.1	**0.4**									0.397	0.490	0.550	0.580							
0.1	**0.5**								0.346	0.356	0.360	0.356	0.346							
0.1	**0.6**								0.392	0.318	0.257	0.207								
0.1	**0.7**								0.423	0.284	0.174									
0.1	**0.8**								0.443	0.254	0.105									
0.1	**0.9**								0.458	0.228	0.048									
0.2	**0.1**									0.417	0.852	1.159	1.350	1.434	1.425	1.331	1.166			
0.2	**0.2**								0.366	0.730	1.000	1.182	1.284	1.312	1.272	1.172	1.018	0.817	0.576	0.301
0.2	**0.3**							0.280	0.560	0.776	0.931	1.029	1.075	1.072	1.024	0.935	0.808	0.647		
0.2	**0.4**							0.477	0.618	0.722	0.793	0.832	0.840	0.819	0.771	0.697				
0.2	**0.5**						0.538	0.582	0.614	0.634	0.640	0.634	0.614	0.582	0.538					
0.2	**0.6**						0.699	0.638	0.583	0.534	0.490	0.448	0.408							
0.2	**0.7**						0.814	0.665	0.540	0.436	0.350	0.280								
0.2	**0.8**						0.898	0.676	0.492	0.342	0.221	0.129								
0.2	**0.9**								0.443	0.254	0.105									
0.3	**0.1**									0.430	0.654	0.808	0.899	0.931	0.911					
0.3	**0.2**							0.280	0.560	0.776	0.931	1.029	1.075	1.072	1.024	0.935	0.808	0.647		
0.3	**0.3**						0.328	0.568	0.759	0.902	1.000	1.055	1.070	1.046	0.986	0.893	0.768	0.614	0.433	0.228
0.3	**0.4**					0.388	0.567	0.711	0.823	0.903	0.953	0.973	0.965	0.929	0.867	0.780	0.668	0.533		
0.3	**0.5**				0.538	0.630	0.706	0.764	0.806	0.832	0.840	0.832	0.806	0.764	0.706	0.630	0.538			
0.3	**0.6**				0.812	0.796	0.780	0.762	0.742	0.719	0.691	0.659	0.621	0.576	0.524					
0.3	**0.7**			1.154	1.024	0.911	0.811	0.725	0.649	0.583	0.524	0.472	0.424	0.379						
0.3	**0.8**						0.814	0.665	0.540	0.436	0.350	0.280								
0.3	**0.9**								0.423	0.284	0.174									
0.4	**0.1**									0.397	0.490	0.550	0.580							
0.4	**0.2**							0.477	0.618	0.722	0.793	0.832	0.840	0.819	0.771	0.697				
0.4	**0.3**					0.388	0.567	0.711	0.823	0.903	0.953	0.973	0.965	0.929	0.867	0.780	0.668	0.533		
0.4	**0.4**				0.432	0.594	0.728	0.835	0.916	0.971	1.000	1.004	0.984	0.940	0.872	0.781	0.668	0.533	0.376	0.198
0.4	**0.5**		0.346	0.490	0.614	0.720	0.806	0.874	0.922	0.950	0.960	0.950	0.922	0.874	0.806	0.720	0.614	0.490	0.346	
0.4	**0.6**	0.543	0.617	0.682	0.737	0.783	0.819	0.844	0.859	0.861	0.852	0.830	0.796	0.748	0.686	0.610	0.519			
0.4	**0.7**				0.812	0.796	0.780	0.762	0.742	0.719	0.691	0.659	0.621	0.576	0.524					
0.4	**0.8**						0.699	0.638	0.583	0.534	0.490	0.448	0.408							
0.4	**0.9**								0.392	0.318	0.257	0.207								
0.5	**0.1**								0.346	0.356	0.360	0.356	0.346							
0.5	**0.2**						0.538	0.582	0.614	0.634	0.640	0.634	0.614	0.582	0.538					
0.5	**0.3**				0.538	0.630	0.706	0.764	0.806	0.832	0.840	0.832	0.806	0.764	0.706	0.630	0.538			
0.5	**0.4**		0.346	0.490	0.614	0.720	0.806	0.874	0.922	0.950	0.960	0.950	0.922	0.874	0.806	0.720	0.614	0.490	0.346	
0.5	**0.5**	0.190	0.360	0.510	0.640	0.750	0.840	0.910	0.960	0.990	1.000	0.990	0.960	0.910	0.840	0.750	0.640	0.510	0.360	0.190
0.5	**0.6**		0.346	0.490	0.614	0.720	0.806	0.874	0.922	0.950	0.960	0.950	0.922	0.874	0.806	0.720	0.614	0.490	0.346	
0.5	**0.7**				0.538	0.630	0.706	0.764	0.806	0.832	0.840	0.832	0.806	0.764	0.706	0.630	0.538			
0.5	**0.8**							0.582	0.614	0.634	0.640	0.634	0.614	0.582	0.538					
0.5	**0.9**								0.346	0.356	0.360	0.356	0.346							
0.6	**0.1**								0.392	0.318	0.257	0.207								
0.6	**0.2**						0.699	0.638	0.583	0.534	0.409	0.448	0.408							
0.6	**0.3**				0.812	0.796	0.780	0.762	0.742	0.719	0.691	0.659	0.621	0.576	0.524					
0.6	**0.4**	0.543	0.617	0.682	0.737	0.783	0.819	0.844	0.859	0.861	0.852	0.830	0.796	0.748	0.686	0.610	0.519			
0.6	**0.5**		0.346	0.490	0.614	0.720	0.806	0.874	0.922	0.950	0.960	0.950	0.922	0.874	0.806	0.720	0.614	0.490	0.346	
0.6	**0.6**				0.432	0.594	0.728	0.835	0.916	0.971	1.000	1.004	0.984	0.940	0.872	0.781	0.668	0.533	0.376	0.198
0.6	**0.7**					0.388	0.567	0.711	0.823	0.903	0.953	0.973	0.965	0.929	0.867	0.780	0.668	0.533		
0.6	**0.8**							0.477	0.618	0.722	0.793	0.832	0.840	0.819	0.771	0.697				
0.6	**0.9**									0.397	0.490	0.550	0.580							
0.7	**0.1**								0.423	0.284	0.174									
0.7	**0.2**						0.814	0.665	0.540	0.436	0.350	0.280								
0.7	**0.3**			1.154	1.024	0.911	0.811	0.725	0.649	0.583	0.524	0.472	0.424	0.379						
0.7	**0.4**				0.812	0.796	0.780	0.762	0.742	0.719	0.691	0.659	0.621	0.576	0.524					
0.7	**0.5**				0.538	0.630	0.706	0.764	0.806	0.832	0.840	0.832	0.806	0.764	0.706	0.630	0.538			
0.7	**0.6**					0.388	0.567	0.711	0.823	0.903	0.953	0.973	0.965	0.929	0.867	0.780	0.668	0.533		
0.7	**0.7**						0.328	0.568	0.759	0.902	1.000	1.055	1.070	1.046	0.986	0.893	0.768	0.614	0.433	0.228
0.7	**0.8**							0.280	0.560	0.776	0.931	1.029	1.075	1.072	1.024	0.935	0.808	0.647		
0.7	**0.9**									0.430	0.654	0.808	0.899	0.931	0.911					
0.8	**0.1**								0.443	0.254	0.105									
0.8	**0.2**						0.898	0.676	0.492	0.342	0.221	0.129								
0.8	**0.3**						0.814	0.665	0.540	0.436	0.350	0.280								
0.8	**0.4**						0.699	0.638	0.583	0.534	0.490	0.448	0.408							
0.8	**0.5**						0.538	0.582	0.614	0.634	0.640	0.634	0.614	0.582	0.538					
0.8	**0.6**							0.477	0.618	0.722	0.793	0.832	0.840	0.819	0.771	0.697				
0.8	**0.7**							0.280	0.560	0.776	0.931	1.029	1.075	1.072	1.024	0.935	0.808	0.647		
0.8	**0.8**								0.366	0.730	1.000	1.182	1.284	1.312	1.272	1.172	1.018	0.817	0.576	0.301
0.8	**0.9**									0.417	0.852	1.159	1.350	1.434	1.425	1.331	1.166			
0.9	**0.1**								0.458	0.228	0.048									
0.9	**0.2**								0.443	0.254	0.105									
0.9	**0.3**								0.423	0.284	0.174									
0.9	**0.4**								0.392	0.318	0.257	0.207								
0.9	**0.5**								0.346	0.356	0.360	0.356	0.346							
0.9	**0.6**									0.397	0.490	0.550	0.580							
0.9	**0.7**									0.430	0.654	0.808	0.899	0.931	0.911					
0.9	**0.8**									0.417	0.852	1.159	1.350	1.434	1.425	1.331	1.166			
0.9	**0.9**									0.169	1.000	1.598	1.984	2.179	2.205	2.083	1.835	1.481	1.043	0.542

Table 2 Values of Q_M for κ_M.

$\pi_{1.}$	$\pi_{.1}$	κ_M -0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	**0.1**									0.169	1.000	1.598	1.984	2.179	2.205	2.083	1.835	1.481	1.043	0.542
0.1	**0.2**									0.464	0.922	1.236	1.419	1.481	1.435	1.292	1.064			
0.1	**0.3**								0.192	0.572	0.840	1.001	1.062	1.030	0.910					
0.1	**0.4**								0.415	0.651	0.793	0.846	0.811	0.692						
0.1	**0.5**								0.608	0.736	0.779	0.736	0.608							
0.1	**0.6**							0.692	0.811	0.846	0.793	0.651	0.415							
0.1	**0.7**						0.910	1.030	1.062	1.001	0.840	0.572	0.192							
0.1	**0.8**				1.064	1.292	1.435	1.481	1.419	1.236	0.922	0.464								
0.1	**0.9**	0.542	1.043	1.481	1.835	2.083	2.205	2.179	1.984	1.598	1.000	0.169								
0.2	**0.1**									0.464	0.922	1.236	1.419	1.481	1.435	1.292	1.064			
0.2	**0.2**								0.366	0.730	1.000	1.182	1.284	1.312	1.272	1.172	1.018	0.817	0.576	0.301
0.2	**0.3**							0.292	0.591	0.820	0.982	1.081	1.121	1.105	1.039	0.925	0.768	0.571		
0.2	**0.4**						0.288	0.540	0.734	0.874	0.960	0.995	0.979	0.916	0.806	0.653	0.456			
0.2	**0.5**						0.566	0.735	0.855	0.928	0.952	0.928	0.855	0.735	0.566					
0.2	**0.6**				0.456	0.653	0.806	0.916	0.979	0.995	0.960	0.874	0.734	0.540	0.288					
0.2	**0.7**			0.571	0.768	0.925	1.039	1.105	1.121	1.081	0.982	0.820	0.591	0.292						
0.2	**0.8**	0.301	0.576	0.817	1.018	1.172	1.272	1.312	1.284	1.182	1.000	0.730	0.366							
0.2	**0.9**				1.064	1.292	1.435	1.481	1.419	1.236	0.922	0.464								
0.3	**0.1**									0.572	0.840	1.001	1.062	1.030	0.910					
0.3	**0.2**							0.292	0.591	0.820	0.982	1.081	1.121	1.105	1.039	0.925	0.768	0.571		
0.3	**0.3**						0.328	0.568	0.759	0.902	1.000	1.055	1.070	1.046	0.986	0.893	0.768	0.614	0.433	0.228
0.3	**0.4**					0.377	0.574	0.733	0.856	0.943	0.996	1.015	1.001	0.956	0.881	0.777	0.644	0.484	0.299	
0.3	**0.5**				0.439	0.608	0.747	0.854	0.931	0.977	0.992	0.977	0.931	0.854	0.747	0.608	0.439			
0.3	**0.6**		0.299	0.484	0.644	0.777	0.881	0.956	1.001	1.015	0.996	0.943	0.856	0.733	0.574	0.377				
0.3	**0.7**	0.228	0.433	0.614	0.768	0.893	0.986	1.046	1.070	1.055	1.000	0.902	0.759	0.568	0.328					
0.3	**0.8**			0.571	0.768	0.925	1.039	1.105	1.121	1.081	0.982	0.820	0.591	0.292						
0.3	**0.9**						0.910	1.030	1.062	1.001	0.840	0.572	0.192							
0.4	**0.1**								0.415	0.651	0.793	0.846	0.811	0.692						
0.4	**0.2**							0.540	0.734	0.874	0.960	0.995	0.979	0.916	0.806	0.653				
0.4	**0.3**					0.377	0.574	0.733	0.856	0.943	0.996	1.015	1.001	0.956	0.881	0.777	0.644	0.484		
0.4	**0.4**				0.432	0.594	0.728	0.835	0.916	0.971	1.000	1.004	0.984	0.940	0.872	0.781	0.668	0.533	0.376	0.198
0.4	**0.5**		0.281	0.449	0.595	0.719	0.820	0.898	0.955	0.988	1.000	0.988	0.955	0.898	0.820	0.719	0.595	0.449	0.281	
0.4	**0.6**	0.198	0.376	0.533	0.668	0.781	0.872	0.940	0.984	1.004	1.000	0.971	0.916	0.835	0.728	0.594	0.432			
0.4	**0.7**		0.299	0.484	0.644	0.777	0.881	0.956	1.001	1.015	0.996	0.943	0.856	0.733	0.574	0.377				
0.4	**0.8**				0.456	0.653	0.806	0.916	0.979	0.995	0.960	0.874	0.734	0.540	0.288					
0.4	**0.9**							0.692	0.811	0.846	0.793	0.651	0.415	0.084						
0.5	**0.1**								0.608	0.736	0.779	0.736	0.608							
0.5	**0.2**						0.566	0.735	0.855	0.928	0.952	0.928	0.855	0.735	0.566					
0.5	**0.3**				0.439	0.608	0.747	0.854	0.931	0.977	0.992	0.977	0.931	0.854	0.747	0.608	0.439			
0.5	**0.4**		0.281	0.449	0.595	0.719	0.820	0.898	0.955	0.988	1.000	0.988	0.955	0.898	0.820	0.719	0.595	0.449	0.281	
0.5	**0.5**	0.190	0.360	0.510	0.640	0.750	0.840	0.910	0.960	0.990	1.000	0.990	0.960	0.910	0.840	0.750	0.640	0.510	0.360	0.190
0.5	**0.6**		0.281	0.449	0.595	0.719	0.820	0.898	0.955	0.988	1.000	0.988	0.955	0.898	0.820	0.719	0.595	0.449	0.281	
0.5	**0.7**				0.439	0.608	0.747	0.854	0.931	0.977	0.992	0.977	0.931	0.854	0.747	0.608	0.439			
0.5	**0.8**						0.566	0.735	0.855	0.928	0.952	0.928	0.855	0.735	0.566					
0.5	**0.9**								0.608	0.736	0.779	0.736	0.608							
0.6	**0.1**							0.692	0.811	0.846	0.793	0.651	0.415							
0.6	**0.2**					0.653	0.806	0.916	0.979	0.995	0.960	0.874	0.734	0.540						
0.6	**0.3**			0.484	0.644	0.777	0.881	0.956	1.001	1.015	0.996	0.943	0.856	0.733	0.574	0.377				
0.6	**0.4**	0.198	0.376	0.533	0.668	0.781	0.872	0.940	0.984	1.004	1.000	0.971	0.916	0.835	0.728	0.594	0.432			
0.6	**0.5**		0.281	0.449	0.595	0.719	0.820	0.898	0.955	0.988	1.000	0.988	0.955	0.898	0.820	0.719	0.595	0.449	0.281	
0.6	**0.6**				0.432	0.594	0.728	0.835	0.916	0.971	1.000	1.004	0.984	0.940	0.872	0.781	0.668	0.533	0.376	0.198
0.6	**0.7**					0.377	0.574	0.733	0.856	0.943	0.996	1.015	1.001	0.956	0.881	0.777	0.644	0.484	0.299	
0.6	**0.8**							0.540	0.734	0.874	0.960	0.995	0.979	0.916	0.806	0.653	0.456			
0.6	**0.9**								0.415	0.651	0.793	0.846	0.811	0.692						
0.7	**0.1**						0.910	1.030	1.062	1.001	0.840	0.572								
0.7	**0.2**			0.571	0.768	0.925	1.039	1.105	1.121	1.081	0.982	0.820	0.591	0.292						
0.7	**0.3**	0.228	0.433	0.614	0.768	0.893	0.986	1.046	1.070	1.055	1.000	0.902	0.759	0.568	0.328					
0.7	**0.4**			0.484	0.644	0.777	0.881	0.956	1.001	1.015	0.996	0.943	0.856	0.733	0.574	0.377				
0.7	**0.5**				0.439	0.608	0.747	0.854	0.931	0.977	0.992	0.977	0.931	0.854	0.747	0.608	0.439			
0.7	**0.6**					0.377	0.574	0.733	0.856	0.943	0.996	1.015	1.001	0.956	0.881	0.777	0.644	0.484		
0.7	**0.7**						0.328	0.568	0.759	0.902	1.000	1.055	1.070	1.046	0.986	0.893	0.768	0.614	0.433	0.228
0.7	**0.8**							0.292	0.591	0.820	0.982	1.081	1.121	1.105	1.039	0.925	0.768	0.571		
0.7	**0.9**								0.192	0.572	0.840	1.001	1.062	1.030	0.910					
0.8	**0.1**				1.064	1.292	1.435	1.481	1.419	1.236	0.922	0.464								
0.8	**0.2**	0.301	0.576	0.817	1.018	1.172	1.272	1.312	1.284	1.182	1.000	0.730	0.366							
0.8	**0.3**			0.571	0.768	0.925	1.039	1.105	1.121	1.081	0.982	0.820	0.591	0.292						
0.8	**0.4**					0.653	0.806	0.916	0.979	0.995	0.960	0.874	0.734	0.540						
0.8	**0.5**						0.566	0.735	0.855	0.928	0.952	0.928	0.855	0.735	0.566					
0.8	**0.6**							0.540	0.734	0.874	0.960	0.995	0.979	0.916	0.806	0.653				
0.8	**0.7**							0.292	0.591	0.820	0.982	1.081	1.121	1.105	1.039	0.925	0.768	0.571		
0.8	**0.8**								0.366	0.730	1.000	1.182	1.284	1.312	1.272	1.172	1.018	0.817	0.576	0.301
0.8	**0.9**									0.464	0.922	1.236	1.419	1.481	1.435	1.292	1.064			
0.9	**0.1**	0.542	1.043	1.481	1.835	2.083	2.205	2.179	1.984	1.598	1.000	0.169								
0.9	**0.2**				1.064	1.292	1.435	1.481	1.419	1.236	0.922	0.464								
0.9	**0.3**						0.910	1.030	1.062	1.001	0.840	0.572								
0.9	**0.4**							0.692	0.811	0.846	0.793	0.651	0.415							
0.9	**0.5**								0.608	0.736	0.779	0.736	0.608							
0.9	**0.6**								0.415	0.651	0.793	0.846	0.811	0.692						
0.9	**0.7**									0.572	0.840	1.001	1.062	1.030	0.910					
0.9	**0.8**									0.464	0.922	1.236	1.419	1.481	1.435	1.292	1.064			
0.9	**0.9**									0.169	1.000	1.598	1.984	2.179	2.205	2.083	1.835	1.481	1.043	0.542

expressed to Professor Dr Bikas K. Sinha for his expert and excellent guidance and his useful comments. The authors also are particularly indebted to the Kasetsart University Research and Development Institute (KURDI), Kasetsart University, Thailand for the financial support which enabled this research to be undertaken. The insightful comments from the referees have been extremely useful for the revision and for this the authors are also grateful.

LITERATURE CITED

Abraira, V. and A. Pérez de Vargas. 1999. Generalization of the kappa coefficient for ordinal categorical data, multiple observers and incomplete designs. **Qüestiió** 23: 561-571.

Banerjee, M., M. Capozzoli, L. McSweeney and D. Sinha. 1999. Beyond kappa: A review of interrater agreement measures. **Can. J. Stat.** 27(1): 3-23.

Cantor, A.B. 1996. Sample-size calculations for Cohen's kappa. **Psychol. Methods** 1(2): 150-153.

Cicchetti, D.V. and A.R. Feinstein. 1990. High agreement but low kappa: II. Resolving the paradoxes. **J. Clin. Epid.** 43: 551-558.

Cohen, J. 1960. A coefficient of agreement for nominal scales. **Educ. Psyc. M.** 20(1): 37-46.

Cohen, J. 1968. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. **Psychol. B.** 70(4): 213-220.

Davies, M. and J.L. Fleiss J.L. 1982. Measuring agreement for multinomial data. **Biometrics** 38: 1047-1051.

Donner, A. and A. Donald. 1988. The statistical analysis ofmultiple binarymeasurements. **J. Clin. Epid.** 41: 899-905.

Everitt, B.S. 1968. Moments of the statistics kappa and weighted kappa. **Br. J. Math. S** 21: 97-103.

Feinstein, A.R. and D.V. Cicchetti. 1990. High agreement but low kappa: I. The problems of two paradoxes. **J. Clin. Epid.** 43: 543-549.

Fleiss, J.L. 1981. **Statistical Methods for Raters and Proportions**. John Wiley and Sons Inc, New York. 352 p.

Fleiss, J.L. and N.W. Chilton. 1983. The measurement of interexaminer agreement on periodontal disease. **J. Period. R.** 18: 601–606.

Fleiss, J.L., J. Cohen and B.S. Everitt. 1969. Large sample standard errors of kappa and weighted kappa. **Psychol. B.** 72(5): 323-327.

Graham, P. and R. Jackson. 1993. The analysis of ordinal agreement data: beyond weighted kappa. **J. Clin. Epid.** 46: 1055-1062.

Gross, S.T. 1986. The kappa coefficient of agreement for multiple observers when the number of subjects is small. **Biometrics** 42: 883-893.

Guggenmoos-Holzmann, I. 1993. How reliable are chance-corrected measures of agreement?. **Stat. Med.** 12: 2191-2205.

Guimarães, M.D.C., H.N. Oliveira, L.N. Campos, C.A. Santos, C.E.R. Gomes, S.B. Oliveira, M.I.F. Freitas, F.A. Acúrcio and C.J. Machado. 2008. Reliability and validity of a questionnaire on vulnerability to sexually transmitted infections among adults with chronic mental illness - PESSOAS Project. **Rev. Bras. Psiquiatr.** 30(1): 55-59.

Hildebrand, D.K, J.D. Laing and H. Rosenthal. 1977. **Analysis of Ordinal Data**. Sage Publications Inc, California. 80 p.

Jelles, F., C.A.M. Van Bennekom, G.F. Lankhorst, C.J.P. Sibbel and L.M. Bouter. 1995. Inter- and intra-rater agreement of the rehabilitation activities profile. **J. Clin. Epid.** 48: 407-416.

Jittavisutthikul, S., A. Thongpan, O. Boodde, W. Shumsing and W. Wajjwalku. 2008. Development of Immunological Test Kit for the Detection of Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) in

Swine. **Kasetsart J. (Nat. Sci.)** 42: 19 - 30.

Landis, J.R. and G.G. Koch. 1977. The measurement of observer agreement for categorical data. **Biometrics** 33: 159-174.

Lantz, C.A. and E. Nebenzahl. 1996. Behavior and interpretation of the kappa statistic: resolution of the two paradoxes. **J. Clin. Epid.** 49(4): 431-434.

Liebetrau, A.M. 1983. **Measures of Association**. Sage Publications Inc, California. 96 p.

Pérez, B., V. Abraira, M. Núñez, P. Boixeda, F. Perez Corral and A. Ledo. 1997. Evaluation of agreement among dermatologists in the assessment of the color of Port Wine Stains and their clearance after treatment with the Flaslamp-Pumped Dye Laser. **Dermatology** 194: 127-130.

Prabhasavat, K. and C. Homgade. 2008. Variation of Hepatic Artery by 3-D Reconstruction MDCT Scan of Liver in Siriraj Hospital. **J. Med. Assoc. Thai** 91(11): 1748-1753.

Schouten, H.J.A. 1986. Nominal scale agreement among observers. **Psychometri.** 51: 453-466.

Sinha, B.K., P. Yimprayoon and M. Tiensuwan. 2006. Cohen’s Kappa Statistic: A Critical Appraisal and Some Modifications. **CalCutta Stat. Assoc.** 58: 151-169.