Listy Biometryczne - Biometrical Letters
Vol. 40 (2003), No. 1, 23-34

Kendall's tau-b type measure of association between difference and sum variables for square contingency tables: application to unaided vision data

Nobuko Miyamoto, Makiko Nishime and Sadao Tomizawa*

Department of Information Sciences, Faculty of Science and Technology,
Tokyo University of Science, Noda City, Chiba 278-8510, Japan
*tomizawa@is.noda.tus.ac.jp

SUMMARY

For the analysis of square contingency tables with the same ordered row and column classifications, Goodman (1985, *Annals of Statistics* **13**, 10-69) considered the diamond (DD) model, which indicates no association between the difference and the sum variables. Tomizawa, Miyamoto and Horiguchi (2000, *Japanese J. Biom.* **20**, 181-191) considered a measure to represent the degree of departure from the DD model, being similar to Goodman and Kruscal's gamma measure. This paper proposes a Kendall's tau-b type measure for representing the degree of departure from the DD model. It is applied to two kinds of unaided distance vision data. For these data, the degrees and the patterns of departure from the DD model are compared.

KEY WORDS: concordance, diamond model, gamma measure, Kendall's tau-b, odds-ratio, quasi-independence.

1. Introduction

For an $R \times R$ square contingency table with the same ordered row and column classifications, let p_{ij} denote the probability that an observation will fall in the ith row and jth column of the table ($i = 1, 2, \dots, R; j = 1, 2, \dots, R$). The diamond (DD) model, which was introduced by Goodman (1985), is defined by

$$p_{ij} = \delta_{i-j}\gamma_{i+j} \quad \text{for} \;\; i = 1, 2, \dots, R; j = 1, 2, \dots, R.$$

Without loss of generality, we can set $\delta_0 = \delta_1 = 1$. [We note that Tomizawa (1994) gave an extension of the DD model (though the detail is omitted)]. The DD model states that there is a structure of quasi-independence between the difference-diagonal classification (i.e., the difference between the row and column classifications) and the sum-diagonal classification (i.e., the sum between the row and column classifications). Consider the $(2R-1) \times (2R-1)$ table of the diamond shape formed by rotating the original $R \times R$ table forty-five degrees so that the $2R-1$ difference-diagonals in the original table now form the entries in the rows of the diamond, and the corresponding $2R-1$ sum-diagonals in the original table form the entries in the columns of the diamond. Denote the row and column variables by X and Y, respectively, for the original $R \times R$ table. Also let $U = X - Y$ and $V = X + Y$. Then the DD model indicates that the probability distribution of (U, V) has a structure of quasi-independence.

Let Δ denote the $(2R-1) \times (2R-1)$ table of the diamond shape formed by rotating the $R \times R$ table based on the observed frequencies forty-five degrees. To illustrate Δ, consider the data in Table 1. Table 1, taken from Stuart (1955), is the cross-classification of the unaided distance vision of 7477 women aged 30-39 employed in Royal Ordnance factories in Britain from 1943 to 1946. Table 2 gives the table Δ. For these data, the DD model hypothesizes that Δ will exhibit the quasi-independence, i.e., no association. The value of likelihood ratio chi-squared statistic G^2 for testing goodness-of-fit of the DD model is 28.88 with $(R-2)^2 = 4$ degrees of freedom. Therefore the DD model fits the data in Table 1 poorly.

Table 1. Unaided distance vision of British women: from Stuart (1955)

Right eye grade (X)		Left eye grade (Y)				
		Best (1)	Second (2)	Third (3)	Worst (4)	Total
Best	(1)	1520	266	124	66	1976
Second	(2)	234	1512	432	78	2256
Third	(3)	117	362	1772	205	2456
Worst	(4)	36	82	179	492	789
Total		1907	2222	2507	841	7477

Let S^* denote a set of cells of the diamond shape in the $(2R-1) \times (2R-1)$ table. Thus,

$$S^* = \{(s,t) | s = i - j, t = i + j \quad \text{for} \quad i = 1, 2, \ldots, R; j = 1, 2, \ldots, R\}.$$

Let p^*_{st} denote the corresponding probability for row value s and column value t, $(s,t) \in S^*$, in the $(2R-1) \times (2R-1)$ table; i.e.,

$$p^*_{st} = p_{\frac{s+t}{2}, \frac{t-s}{2}} \quad \text{for} \quad (s,t) \in S^*.$$

Using the odds-ratios defined for the $(2R-1) \times (2R-1)$ table, the DD model may

Table 2. Table Δ constructed from Table 1

Right eye grade minus left eye grade (U)	Right eye grade plus left eye grade (V) 2	3	4	5	6	7	8
-3	*	*	*	66	*	*	*
-2	*	*	124	*	78	*	*
-1	*	266	*	432	*	205	*
0	1520	*	1512	*	1772	*	492
1	*	234	*	362	*	179	*
2	*	*	117	*	82	*	*
3	*	*	*	36	*	*	*

be expressed as

$$\theta^*_{(s<k;t<l)} = 1 \quad \text{for} \quad (s,t),(s,l),(k,t),(k,l) \in S^*, s<k, t<l, \tag{1.1}$$

where

$$\theta^*_{(s<k;t<l)} = (p^*_{st}p^*_{kl})/(p^*_{sl}p^*_{kt}).$$

Note that $\theta^*_{(s<k;t<l)}$ is the odds-ratio which can be defined for rows $s<k$ and columns $t<l$ in the $(2R-1)\times(2R-1)$ table.

We also note that DD model does not impose the restriction on four corner cells of diamond shape, i.e., on cells $(0,2)$, $(0,2R)$, $(1-R,R+1)$, and $(R-1,R+1)$ in S^* for the $(2R-1)\times(2R-1)$ table. Therefore we shall denote S^* except these four cells by D^*. Thus the DD model also may be expressed as

$$\theta^*_{(s<k;t<l)} = 1 \quad \text{for} \quad (s,t),(s,l),(k,t),(k,l) \in D^*, s<k, t<l.$$

When the DD model fits the table data poorly, we are interested in measuring the degree of departure from the DD model. Tomizawa, Miyamoto and Horiguchi (2000) proposed the measure ϕ which represents the degree of departure from the DD model; see Appendix for measure ϕ. The measure ϕ is based on Goodman and Kruskal's (1954) gamma measure γ.

The measure γ is defined by

$$\gamma = \frac{\Pi_c - \Pi_d}{\Pi_c + \Pi_d},$$

where

$$\Pi_c = 2\sum_{s<k}\sum_{t<l} p_{st}p_{kl} \quad , \quad \Pi_d = 2\sum_{s<k}\sum_{t<l} p_{sl}p_{kt} \, .$$

So, the γ measure represents the association (between the row variable X and the column variable Y) for the cross-classification of ordinal variables (see Agresti, 1984, pp.159-160). Also, Kendall's tau-b measure, which was introduced by Kendall (1945), represents the association for the cross-classification of ordinal variables. The Kendall's tau-b measure is defined by

$$\tau_b = \frac{\Pi_c - \Pi_d}{\sqrt{\left[1 - \sum_{s=1}^{R} p_{s+}^2\right]\left[1 - \sum_{t=1}^{R} p_{+t}^2\right]}},$$

where

$$p_{s+} = \sum_{j=1}^{R} p_{sj} \quad , \quad p_{+t} = \sum_{i=1}^{R} p_{it} \, .$$

Since the Kendall's tau-b measure includes pairs tied on both X and Y in the denominator (Agresti, 1984, p.162), the denominator of the Kendall's tau-b measure is at least as large as the denominator of the gamma measure, i.e., $|\tau_b| \leq |\gamma|$, with $-1 \leq \tau_b \leq 1$ and $-1 \leq \gamma \leq 1$.

The purpose of this paper is to propose a measure which is based on Kendall's tau-b measure and represents the degree of departure from the DD model (i.e., represents the association between the difference and sum variables). The measure would be useful for comparing the degrees and the patterns of departure from the DD model for the unaided vision data (in Section 4). We shall consider the Kendall's tau-b type measure of departure from the DD model in Section 2.

2. Kendall's tau-b type measure of departure from DD model

Let

$$\Gamma_c = 2 \sum\sum_{(s<k;t<l)\in D^*} p_{st}^* p_{kl}^* \, ,$$

where $(s < k; t < l) \in D^*$ means $(s,t), (s,l), (k,t), (k,l) \in D^*$ (being four cells so that it is possible to define the odds-ratios), defined by rows $s < k$ and columns $t < l$ in the $(2R-1) \times (2R-1)$ table of the diamond shape. Note that $2p_{st}^* p_{kl}^*$ is the (conditional) probability of concordance for an arbitrary pair of observations on condition that the pair falls in one or two of possible cells $(s,t), (s,l), (k,t), (k,l) \in D^*$ defined for rows $s < k$ and columns $t < l$ in the $(2R-1) \times (2R-1)$ table; namely, for a pair of observations which will fall in one or two of these four cells, $2p_{st}^* p_{kl}^*$ is the (conditional) probability that the member that ranks higher on variable U also ranks higher on variable V.

Also let

$$\Gamma_d = 2 \sum\sum_{(s<k;t<l)\in D^*} p^*_{sl}p^*_{kt} .$$

Note that $2p^*_{sl}p^*_{kt}$ is the (conditional) probability of discordance for an arbitrary pair of observations on condition that the pair falls in one or two of possible cells $(s,t),(s,l),(k,t),(k,l) \in D^*$ defined for rows $s < k$ and columns $t < l$ in the $(2R-1) \times (2R-1)$ table; namely, for a pair of observations which will fall in one or two of these four cells, $2p^*_{sl}p^*_{kt}$ is the (conditional) probability that the member that ranks higher on variable U ranks lower on variable V. In addition, we note that the DD model also indicates that for arbitrary possible cells $(s,t),(s,l),(k,t),(k,l) \in D^*$ defined for rows $s < k$ and columns $t < l$ in the $(2R-1) \times (2R-1)$ table, $p^*_{st}p^*_{kl}$ equals $p^*_{sl}p^*_{kt}$ (from (1.1)); namely, when a pair of observations will fall in one or two of such four cells, the (conditional) probability that the member that ranks higher on variable U also ranks higher on variable V is equal to the (conditional) probability that it ranks lower on variable V.

Let

$$\Gamma_x = 2 \sum\sum_{\substack{(s,t),(s,l)\in D^* \\ t<l}} p^*_{st}p^*_{sl} .$$

We note that Γ_x indicates the probability that a pair for observations unties on variable V but ties on variable U.

Similarly, let

$$\Gamma_y = 2 \sum\sum_{\substack{(s,t),(k,t)\in D^* \\ s<k}} p^*_{st}p^*_{kt} .$$

We note that Γ_y indicates the probability that a pair for observations unties on variable U but ties on variable V.

We shall consider a measure, which represents the degree of departure from the DD model, as follows:

$$\phi_b = \frac{\Gamma_c - \Gamma_d}{\sqrt{(\Gamma_c + \Gamma_d + \Gamma_y)(\Gamma_c + \Gamma_d + \Gamma_x)}} .$$

This measure must lie between -1 and 1. Also (i) $\phi_b = 1$ if and only if $\Gamma_d = \Gamma_x = \Gamma_y = 0$; namely, for any pair of observations (U_a, V_a) and (U_b, V_b) (which falls in one or two of four cells so that it is possible to define the odds-ratios), if $U_a < U_b$, then $V_a < V_b$, and if $V_a < V_b$, then $U_a < U_b$; and (ii) $\phi_b = -1$ if and only if $\Gamma_c = \Gamma_x = \Gamma_y = 0$; namely, for any pair of them, if $U_a < U_b$, then $V_a > V_b$, and if $V_a < V_b$, then $U_a > U_b$. We note that if the DD model holds, then $\phi_b = 0$, but the converse of this does not hold.

For the pair of observations (U_a, V_a) and (U_b, V_b) (which falls in one or two of four cells so that it is possible to define the odds-ratios), let

$$U_{ab} = \mathrm{sign}(U_a - U_b) = \begin{cases} -1 & \text{if} \quad U_a < U_b \ , \\ 0 & \text{if} \quad U_a = U_b \ , \\ 1 & \text{if} \quad U_a > U_b \ , \end{cases}$$

$$V_{ab} = \mathrm{sign}(V_a - V_b) = \begin{cases} -1 & \text{if} \quad V_a < V_b \ , \\ 0 & \text{if} \quad V_a = V_b \ , \\ 1 & \text{if} \quad V_a > V_b \ . \end{cases}$$

The sign scores $\{U_{ab}\}$ are indicator variables that index whether U_a is greater than or less than U_b, and similarly for $\{V_{ab}\}$. Note that $U_{ab} = -U_{ba}$ and $V_{ab} = -V_{ba}$. The product $U_{ab}V_{ab}$ equals 1 for a concordant pair and -1 for a discordant pair. The square U_{ab}^2 equals 1 for a pair untied on U and equals 0 for a pair tied on U. Similarly, V_{ab}^2 equals 1 for a pair untied on V and equals 0 for a pair tied on V.

Let n_{ij} denote the observed frequency in the ith row and jth column of the original $R \times R$ table $(i = 1, 2, \ldots, R; j = 1, 2, \ldots, R)$. Also let n_{st}^* denote the corresponding observed frequency of row value s and column value t, $(s,t) \in S^*$, in the $(2R-1) \times (2R-1)$ table of the diamond shape. The sample version of ϕ_b, i.e., $\hat{\phi}_b$, is given by ϕ_b with $\{p_{st}^*\}$ replaced by $\{\hat{p}_{st}^*\}$, where $\hat{p}_{st}^* = n_{st}^*/n$ for $(s,t) \in S^*$ and $n = \sum\sum_{(s,t)\in S^*} n_{st}^* = \sum_{i=1}^{R}\sum_{j=1}^{R} n_{ij}$.

For an arbitrary pair of observations (U_a, V_a) and (U_b, V_b) (which falls in one or two of four cells so that it is possible to define the odds-ratios), we see

$$\sum\sum_{a \neq b} U_{ab}V_{ab} = 2n^2(\hat{\Gamma}_c - \hat{\Gamma}_d) \ ,$$

$$\sum\sum_{a \neq b} U_{ab}^2 = 2n^2(\hat{\Gamma}_c + \hat{\Gamma}_d + \hat{\Gamma}_y) \ ,$$

$$\sum\sum_{a \neq b} V_{ab}^2 = 2n^2(\hat{\Gamma}_c + \hat{\Gamma}_d + \hat{\Gamma}_x) \ ,$$

$$\sum\sum_{a \neq b} U_{ab} = \sum\sum_{a \neq b} V_{ab} = 0 \ ,$$

where $\hat{\Gamma}_c$, $\hat{\Gamma}_d$, $\hat{\Gamma}_x$ and $\hat{\Gamma}_y$ denote the Γ_c, Γ_d, Γ_x and Γ_y with $\{p_{st}^*\}$ replaced by $\{\hat{p}_{st}^*\}$. Each pair is used twice in these sums so that $\sum\sum U_{ab} = \sum\sum V_{ab} = 0$ due to the relationships $U_{ab} = -U_{ba}$ and $V_{ab} = -V_{ba}$. The sample correlation between the $\{U_{ab}\}$

and $\{V_{ab}\}$ is therefore

$$\frac{\sum\sum_{a\neq b} U_{ab}V_{ab}}{\sqrt{(\sum\sum_{a\neq b} U_{ab}^2)(\sum\sum_{a\neq b} V_{ab}^2)}} = \hat{\phi}_b \ .$$

Thus the measure ϕ_b is a special case of correlation defined using pair scores.

We note that the measure ϕ_b is somewhat similar to the Kendall's tau-b measure τ_b. The measure ϕ_b represents the association (between the difference-variable U and the sum-variable V) for the Δ table, though the measure τ_b represents the association (between the row variable X and the column variable Y) for the original table.

3. Approximate confidence interval for measure

Assuming that a multinomial distribution applies to the $R\times R$ table, we shall consider an approximate standard error and large-sample confidence interval for ϕ_b using the delta method, descriptions of which are given by Bishop, Fienberg and Holland (1975, Sec.14.6) and Agresti (1984, p.185, Appendix C). Let $\delta_x = (\Gamma_c+\Gamma_d+\Gamma_x)^{\frac{1}{2}}$, $\delta_y = (\Gamma_c+\Gamma_d+\Gamma_y)^{\frac{1}{2}}$ and $\delta = \delta_x\delta_y$. Using the delta method, $\sqrt{n}(\hat{\phi}_b-\phi_b)$ has asymptotically (as $n\to\infty$) a normal distribution with mean zero and variance

$$\begin{aligned}\sigma^2[\phi_b] &= \frac{1}{4\delta^4}\sum\sum_{(s,t)\in D^*} p_{st}^*\{2\delta(p_{st}^{*(c)}-p_{st}^{*(d)})-\phi_b\delta_x^2(p_{st}^{*(c)}+p_{st}^{*(d)}+p_{st}^{*(y)})-\\ &\quad \phi_b\delta_y^2(p_{st}^{*(c)}+p_{st}^{*(d)}+p_{st}^{*(x)})\}^2 \ ,\end{aligned}$$

where

$$p_{st}^{*(c)} = \sum\sum_{(a<s;b<t)\in D^*} p_{ab}^* + \sum\sum_{(s<a;t<b)\in D^*} p_{ab}^* \ ,$$

$$p_{st}^{*(d)} = \sum\sum_{(a<s;t<b)\in D^*} p_{ab}^* + \sum\sum_{(s<a;b<t)\in D^*} p_{ab}^* \ ,$$

$$p_{st}^{*(x)} = \sum\sum_{\substack{(s,b)\in D^*\\ b\neq t}} p_{sb}^* \ ,$$

$$p_{st}^{*(y)} = \sum\sum_{\substack{(a,t)\in D^*\\ a\neq s}} p_{at}^* \ .$$

Let $\hat{\sigma}^2[\phi_b]$ denote $\sigma^2[\phi_b]$ with $\{p_{st}^*\}$ replaced by $\{\hat{p}_{st}^*\}$. Then $\hat{\sigma}[\phi_b]/\sqrt{n}$ is an estimated approximate standard error for ϕ_b, and $\hat{\phi}_b \pm z_{p/2}\hat{\sigma}[\phi_b]/\sqrt{n}$ is an approximate $100(1-p)$ percent confidence interval for ϕ_b, where $z_{p/2}$ is the percentage point from the standard normal distribution corresponding to a two-tail probability equal to p.

4. Analysis of vision data

Table 3 taken from Tomizawa (1984) and Tomizawa, Miyamoto and Horiguchi (2000) is the data on unaided distance vision of 4746 students aged 18 to about 25 including women of about 10% in Faculty of Science and Technology, Science University of Tokyo in Japan examined in April, 1982. Table 4 gives the table Δ constructed from Table 3. The value of likelihood ratio chi-squared statistic G^2 for testing goodness of fit of the DD model is 107.80 for Table 3 with both $(R-2)^2 = 4$ degrees of freedom. (The value of G^2 for the data in Table 1 is described in Section 1). Therefore the DD model fits each of Tables 1 and 3 poorly.

Table 3. Unaided distance vision of students in Japan: from Tomizawa (1984)

Right eye grade (X)	Left eye grade (Y)				
	Best (1)	Second (2)	Third (3)	Worst (4)	Total
Best (1)	1291	130	40	22	1483
Second (2)	149	221	114	23	507
Third (3)	64	124	660	185	1033
Worst (4)	20	25	249	1429	1723
Total	1524	500	1063	1659	4746

Table 4. Table Δ constructed from Table 3

Right eye grade minus left eye grade (U)	Right eye grade plus left eye grade (V)						
	2	3	4	5	6	7	8
-3	*	*	*	22	*	*	*
-2	*	*	40	*	23	*	*
-1	*	130	*	114	*	185	*
0	1291	*	221	*	660	*	1429
1	*	149	*	124	*	249	*
2	*	*	64	*	25	*	*
3	*	*	*	20	*	*	*

We shall apply the measure ϕ_b to the data in Tables 1 and 3. For the data in Table 1, the estimated value of ϕ_b is $\hat{\phi}_b = 0.0059$. The estimated approximate standard error of $\hat{\phi}_b$ is 0.0002, and an approximate 95% confidence interval for ϕ_b is (0.0056, 0.0062). Also for the data in Table 3, the estimated value of ϕ_b is $\hat{\phi}_b = -0.0149$. The estimated approximate standard error of $\hat{\phi}_b$ is 0.0004, and an approximate 95% confidence interval for ϕ_b is $(-0.0156, -0.0142)$. Since both of these confidence intervals do not include zero, these would indicate that there is not the structure of quasi-independence in the Δ tables of Tables 2 and 4 (i.e., the DD model does not hold).

Also, we see that $|\hat{\phi}_b| = 0.0059$ for Table 1 and $|\hat{\phi}_b| = 0.0149$ for Table 3. These may indicate that the magnitude of departure from the DD model is greater for Table 3 than for Table 1.

For the data in Table 1, we shall define a woman's total vision of both eyes by the sum of the right eye grade X and the left eye grade Y (i.e., the total vision for a woman that fell in a cell (i,j) in Table 1 is $i+j$). Then the sum-variable V indicates a woman's total vision. Also we shall define a woman's unbalanced grade of both eyes by the right eye grade X minus the left eye grade Y (i.e., the unbalanced grade of both eyes for a woman that fell in a cell (i,j) in Table 1 is $i-j$). Then the difference-variable U indicates a woman's unbalanced grade of both eyes. [Note that a woman's right eye becomes worse than her left eye as $|U|$ with $U > 0$ increases, and her right eye becomes better than her left eye as $|U|$ with $U < 0$ increases.] The DD model indicates that there is no association between the unbalanced grade of both eyes, U, and the total vision of both eyes, V. (Similarly, U and V can be interpreted for the data in Table 3).

For the data in Table 1 (or in Table 2) and Table 3 (or in Table 4), the DD model fits poorly. Therefore, for these data there is an association between the unbalanced grade of both eyes and the total vision of both eyes. In addition, we can infer a type of association by seeing whether the $\hat{\phi}_b$ is positive or negative. In fact, for the data in Table 1 (or in Table 2), the values of the confidence interval for ϕ_b are positive. Thus, for these data, Γ_c is estimated to be greater than Γ_d; namely, the sum of (conditional) probability of concordance for a randomly selected pair of women (on condition that the pair falls in one or two of four cells so that it is possible to define the odds-ratio in Table 2) is estimated to be greater than the sum of (conditional) probability of discordance for the pair of women. So, roughly speaking, for such a pair of women, the member that ranks higher on the unbalanced grade of both eyes (U) may tend to rank higher (rather than lower) on the total vision of both eyes (V). On the other hand, for the data in Table 3 (or Table 4), the values of the confidence interval for ϕ_b are negative. Thus for these data, Γ_c is estimated to be less than Γ_d; namely, the sum of (conditional) probability of concordance for a randomly selected pair of students is estimated to be less than the sum of (conditional) probability of discordance for the pair of students. So, roughly speaking, for such a pair of students, the member that ranks higher on the unbalanced grade of both eyes (U) may tend to rank lower (rather than higher) on the total vision of both eyes (V).

5. Concluding remarks

Both the Kendall's tau-b type measure ϕ_b and Tomizawa-Miyamoto-Horiguchi measure ϕ (in Appendix) range between -1 and 1 and represent the degree of departure from the DD model.

We point out that the measure ϕ_b counts the numbers of pairs tied on both U and V variables in the denominator though the measure ϕ does not count them. Consider the artificial data for the Δ table in Table 5. Then the estimated values of ϕ_b and ϕ are $\hat{\phi}_b = 0.676$ and $\hat{\phi} = 1.000$. So, $\hat{\phi}_b$ is less than $\hat{\phi}$ which takes the maximum value. For the data in Table 5, we can see that the estimated probability of concordance is $\hat{\Gamma}_c = 0.16$ (> 0), but that of discordance is $\hat{\Gamma}_d = 0.00$ with $\hat{\Gamma}_x = 0.04$ (> 0) and $\hat{\Gamma}_y = 0.12$ (> 0); thus there is a pair of observations (U_a, V_a) and (U_b, V_b) (which falls in one or two of four cells so that it is possible to define the odds-ratios) such that $U_a < U_b$ and $V_a = V_b$ (or $V_a < V_b$ and $U_a = U_b$). Namely, for these data, there is not a pair of observations with discordance, however, there are some pairs of observations with concordance, and also there are some pairs of observations such that the pair unties on variable U but ties on variable V (or the pair unties on V but ties on U). Thus, there is not always concordance for any pair of observations such that the member that ranks higher on variable U also ranks higher on variable V. Therefore, for the data in Table 5, it seems natural to consider that the value of measure of departure from the DD model should be less than 1 which indicates the maximum departure. Therefore, the measure ϕ_b rather than the measure ϕ would be suitable for

Table 5. The artificial data for Δ table

				(V)			
(U)	2	3	4	5	6	7	8
-3	*	*	*	0	*	*	*
-2	*	*	40	*	0	*	*
-1	*	0	*	40	*	0	*
0	0	*	0	*	20	*	0
1	*	0	*	40	*	20	*
2	*	*	0	*	40	*	*
3	*	*	*	0	*	*	*

measuring the degree of departure from the DD model, i.e., for measuring the degree of the association between the difference-variable U and the sum-variable V.

In more detail, we note that $\phi_b = 1$ if and only if, for the pair of observations (U_a, V_a) and (U_b, V_b) (in possible cells as above), if $U_a < U_b$, then $V_a < V_b$, and if $V_a < V_b$, then $U_a < U_b$ [though $\phi = 1$ (for ϕ in Appendix) if and only if $\Gamma_d = 0$, namely, if $U_a < U_b$, then $V_a < V_b$ or $V_a = V_b$, and if $V_a < V_b$, then $U_a < U_b$ or $U_a = U_b$]. So, we point out that it would be natural that the case of maximum departure from equality of concordance and discordance should be defined by the case of $\phi_b = 1$ (of which the condition is stronger than that of $\phi = 1$). So, the measure ϕ_b rather than the measure ϕ would be suitable.

Appendix

The gamma type measure of departure from the DD model, which was considered by Tomizawa, Miyamoto and Horiguchi (2000), is defined by

$$\phi = \frac{\Gamma_c - \Gamma_d}{\Gamma_c + \Gamma_d},$$

where

$$\Gamma_c = 2 \sum\sum_{(s<k;t<l)\in D^*} p^*_{st} p^*_{kl},$$

$$\Gamma_d = 2 \sum\sum_{(s<k;t<l)\in D^*} p^*_{sl} p^*_{kt},$$

and where $(s < k; t < l) \in D^*$ means $(s, t), (s, l), (k, t), (k, l) \in D^*$(being four cells so that it is possible to define the odds-ratios), defined by rows $s < k$ and columns $t < l$ in the $(2R - 1) \times (2R - 1)$ table of the diamond shape.

References

Agresti, A. (1984). *Analysis of Ordinal Categorical Data.* John Wiley, New York.

Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. (1975). *Discrete Multivariate Analysis: Theory and Practice.* Cambridge, Massachusetts: The MIT Press.

Goodman, L.A. (1985). The analysis of cross-classified data having ordered and/or unordered categories: association models, correlation models, and asymmetry models for contingency tables with or without missing data. *Annals of Statistics* **13**, 10-69.

Goodman, L.A., and Kruskal, W.H. (1954). Measures of association for cross classification. *Journal of the American Statistical Association* **49**, 732-764.

Kendall, M. G. (1945). The treatment of ties in rank problems. *Biometrika* **33**, 239-251.

Stuart, A. (1955). A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika* **42**, 412-416.

Tomizawa, S. (1984). Three kinds of decompositions for the conditional symmetry model in a square contingency table. *Journal of the Japan Statistical Society* **14**, 35-42.

Tomizawa, S. (1994). Uniform association diamond model for square contingency tables with ordered categories. *Journal of the Japan Statistical Society* **24**, 83-88.

Tomizawa, S., Miyamoto, N., and Horiguchi, H. (2000). Measure of departure from diamond model and its application to unaided vision data. *Japanese Journal of Biometrics* **20**, 181-191.

Received 10 August 2003

Miara typu tau-b Kendalla związku pomiędzy zmiennymi w kwadratowej tablicy kontyngencji: zastosowanie do danych opisujących jakość widzenia

STRESZCZENIE

Dla analizy kwadratowych tablic kontyngencji utworzonych przez tę samą, uporządkowaną klasyfikację wierszową i kolumnową, Goodman (1985) rozważał model romboidalny, który opisuje brak związku pomiędzy zmiennymi utworzonymi poprzez odejmowanie i dodawanie wartości oryginalnych zmiennych. Tomizawa i in. (2000) rozważali pewną miarę odchylenia od tego modelu podobną do miary gamma. W pracy proponuje się miarę odchylenia podobną do miary tau-b Kendalla. Przedstawiony przykład dotyczy danych opisujących jakość widzenia prawym i lewym okiem.

Słowa kluczowe: zgodność, model romboidalny, miara gamma, tau-b Kendalla, iloraz szans, quasi-niezależność