



Partial Asymmetry Measures for Square Contingency Tables

Takuma Ishihara ^{1,*} , Kouji Yamamoto ², Kouji Tahata ³  and Sadao Tomizawa ⁴¹ Innovative and Clinical Research Promotion Center, Gifu University Hospital, Gifu City 501-1194, Gifu, Japan² Department of Biostatistics, Yokohama City University School of Medicine, Yokohama City 236-0004, Kanagawa, Japan³ Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda City 278-8510, Chiba, Japan⁴ Department of Information Sciences, Faculty of Information Sciences, Meisei University, Hino City 191-0042, Tokyo, Japan

* Correspondence: tishiha@gifu-u.ac.jp; Tel.: +81-58-230-6655

Abstract: In square contingency table analysis, we consider a partial measure that represents the degree of departure from symmetry for each of several pairs. It may be useful to pool the values of the measure into a single summary measure of partial asymmetry. We show that the estimator of partial measures is asymptotically mutually independent for a large sample size. The present paper proposes a symmetry measure in the class of weighted averages that is different from previous studies. The proposed measure is an approximation of the measure in the class of weighted averages that has the smallest variance.

Keywords: categorical data; Kullback–Leibler information; minimum variance; nominal category; Shannon entropy



Citation: Ishihara, T.; Yamamoto, K.; Tahata, K.; Tomizawa, S. Partial Asymmetry Measures for Square Contingency Tables. *Symmetry* **2022**, *14*, 1936. <https://doi.org/10.3390/sym14091936>

Academic Editor: Jinyu Li

Received: 19 August 2022

Accepted: 14 September 2022

Published: 17 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In categorical data analysis, contingency tables are a basic tool used to examine the relationship between row and column categories. For example, the Pearson X^2 statistic is commonly used to test the null hypothesis of statistical independence (Agresti [1] [p. 75]). When statistical independence is rejected, we are interested in describing the association between the row and column categories. Summary measures of association have been proposed, such as the *Cramér V*, *gamma*, and *uncertainty coefficient*. For details, see for instance Agresti [1] [Sec. 2.4] and Bishop et al. [2] [Sec. 11.3]. Additionally, the recent development of association measures is described, for example, in Beh et al. [3], Lombardo [4], Wei and Kim [5,6], Zhang et al. [7], and Wei et al. [8].

Contingency tables with the same row and column classifications are called *square* contingency tables. These tables are used for unaided distance vision data, social mobility data, and longitudinal data in biomedical research. The analysis of square contingency tables considers the issue of symmetry rather than independence because it is not sensible to treat these data as independent.

Bowker [9] introduced the simple symmetry model and proposed a test for the hypothesis of symmetry. When the symmetry model fits the given data poorly, we are interested in measuring the degree of departure from symmetry. Tomizawa [10] proposed a measure that represents the degree of departure from symmetry expressed using the Shannon entropy or Kullback–Leibler information. In the real world, the Shannon entropy is widely applied as a measure of complexity, for example in Fernandes and Araújo [11]. The measure lies between 0 and 1, and its value equals 0 if and only if the symmetry model holds. Additionally, the degree of departure from symmetry increases as the value of the measure increases.

In the present paper, we propose a measure that represents the degree of departure from symmetry using a different approach. We also consider a partial measure that represents the degree of departure from symmetry for each of several pairs. If the asymmetry

appears to be similar in the various pairs, it may be useful to pool the values of the measure into a single summary measure of partial asymmetry. In an analogous manner to Agresti [12] [p. 170], we consider taking a weighted average of the sample values as a summary measure. The properties of the proposed measure are given, and it has a characteristic that is different from that of Tomizawa’s measure.

The rest of this paper is organized as follows. Section 2 describes the background of this study by reviewing previous research. Section 3 proposes the new measure that represents the degree of departure from symmetry. Section 4 shows some numerical examples and discusses the difference between the estimate of Φ_T and the estimate of the proposed measure. Section 5 gives an example involving the cross-classification of mothers’ and fathers’ birth orders. Section 6 contains concluding remarks.

2. Review of Previous Research

Consider an $r \times r$ square contingency table having the same row and column classifications. Let p_{ij} denote the probability that an observation will fall in the (i, j) th cell of the table ($i = 1, \dots, r; j = 1, \dots, r$). The simple symmetry model introduced by Bowker [9] is defined by

$$p_{ij} = p_{ji} \quad (i \neq j). \tag{1}$$

This model indicates the symmetry structure with respect to the cell probabilities. Bowker [9] proposed a test for the hypothesis of symmetry.

When the symmetry model does not hold for a given dataset, we are interested in evaluating the degree of departure from symmetry. Assuming $p_{ij} + p_{ji}$ is not equal to zero for $i < j$, the measure is defined as

$$\Phi_T = \frac{1}{\delta \log 2} \sum_{i \neq j} p_{ij} \log \frac{2p_{ij}}{p_{ij} + p_{ji}}, \tag{2}$$

where $\delta = \sum_{i \neq j} p_{ij}$. The measure Φ_T has three properties: (i) $0 \leq \Phi_T \leq 1$; (ii) the table has a symmetrical structure if and only if $\Phi_T = 0$; (iii) there is a structure for which either $p_{ij} = 0$ or $p_{ji} = 0$ for $i \neq j$ if and only if $\Phi_T = 1$.

Let $\pi_{ij} = p_{ij} / (p_{ij} + p_{ji})$ for $i = 1, \dots, r; j = 1, \dots, r; i \neq j$. The conditional probability that an observation falls in cell (i, j) or (j, i) in the table is π_{ij} . It should be noted that the symmetry model can be expressed as

$$\pi_{ij} = \pi_{ji} \quad \left(= \frac{1}{2} \right) \quad (i < j). \tag{3}$$

The measure Φ_T can be expressed as

$$\Phi_T = \sum_{i < j} \left(\frac{p_{ij} + p_{ji}}{\delta} \right) \phi_{ij}, \tag{4}$$

where

$$\phi_{ij} = \frac{1}{\log 2} \left(\pi_{ij} \log \frac{\pi_{ij}}{1/2} + \pi_{ji} \log \frac{\pi_{ji}}{1/2} \right). \tag{5}$$

It should be noted that ϕ_{ij} is the normalized Kullback–Leibler information between (π_{ij}, π_{ji}) and $(1/2, 1/2)$. That is, the measure Φ_T is the weighted average of ϕ_{ij} .

We review ϕ_{ij} in Φ_T . The partial measure ϕ_{ij} represents the degree of departure from symmetry for a pair of symmetric cells because: (i) $0 \leq \phi_{ij} \leq 1$; (ii) there is a symmetrical structure for the pair of (i, j) and (j, i) cells if and only if $\phi_{ij} = 0$; (iii) there is a structure for which either $p_{ij} = 0$ or $p_{ji} = 0$ for the pair of (i, j) and (j, i) cells if and only if $\phi_{ij} = 1$. That is, the measure ϕ_{ij} expresses partial asymmetry.

3. The Proposed Measure

Let n_{ij} denote the observed frequency in the (i, j) th cell of the table ($i = 1, \dots, r; j = 1, \dots, r$). We assume that $\{n_{ij}\}$ have a multinomial distribution:

$$\frac{n!}{\prod_{i=1}^r \prod_{j=1}^r n_{ij}!} \prod_{i=1}^r \prod_{j=1}^r p_{ij}^{n_{ij}}, \tag{6}$$

where $n = \sum_{i=1}^r \sum_{j=1}^r n_{ij}$. Let \hat{p}' be the $1 \times r^2$ vector

$$\hat{p}' = (\hat{p}'_{(12)}, \hat{p}'_{(13)}, \dots, \hat{p}'_{(r-1,r)}, \hat{p}_{11}, \dots, \hat{p}_{rr}), \tag{7}$$

where

$$\hat{p}'_{(ij)} = (\hat{p}_{ij}, \hat{p}_{ji}), \quad \hat{p}_{ij} = \frac{n_{ij}}{n}, \quad \hat{p}_{ii} = \frac{n_{ii}}{n}. \tag{8}$$

Furthermore, let us define the vector p in terms of p_{ij} s in the same way as \hat{p} . Let $\hat{\phi}_{ij}$ denote the sample version of ϕ_{ij} . Namely, the estimated ϕ_{ij} is given as

$$\hat{\phi}_{ij} = \frac{1}{\log 2} \left(\hat{\pi}_{ij} \log \frac{\hat{\pi}_{ij}}{1/2} + \hat{\pi}_{ji} \log \frac{\hat{\pi}_{ji}}{1/2} \right), \tag{9}$$

where $\hat{\pi}_{ij} = \hat{p}_{ij} / (\hat{p}_{ij} + \hat{p}_{ji})$ and $\hat{\pi}_{ji} = \hat{p}_{ji} / (\hat{p}_{ij} + \hat{p}_{ji})$. Let $\hat{\phi}$ be the $r(r-1)/2 \times 1$ vector:

$$\hat{\phi} = (\hat{\phi}_{12}, \hat{\phi}_{13}, \dots, \hat{\phi}_{r-1,r})', \tag{10}$$

and we define the vector ϕ in terms of ϕ_{ij} s in a similar manner to $\hat{\phi}$. From Appendix A, $\hat{\phi}$ is asymptotically distributed as normal with mean ϕ and covariance matrix

$$\sigma^2[\hat{\phi}] = \begin{pmatrix} \sigma_{12}^2 & 0 & \dots & 0 \\ 0 & \sigma_{13}^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \sigma_{r-1,r}^2 \end{pmatrix}, \tag{11}$$

where

$$\sigma_{ij}^2 = \frac{\pi_{ij}\pi_{ji}}{n(p_{ij} + p_{ji})} \left(\frac{1}{\log 2} (\log \pi_{ij} - \log \pi_{ji}) \right)^2 \quad (i < j). \tag{12}$$

It should be noted that the set $\hat{\phi}_{12}, \dots, \hat{\phi}_{r-1,r}$ is asymptotically mutually independent for large n .

We consider the weighted average of $\{\hat{\phi}_{ij}\}$, that is

$$\Phi = \sum_{i < j} \sum w_{ij} \hat{\phi}_{ij}, \tag{13}$$

where the weights $\{w_{ij}\}$ satisfy all $w_{ij} > 0$ and $\sum_{i < j} \sum w_{ij} = 1$. We see from Appendix A that $\hat{\phi}_{ij}$ ($i < j$) is asymptotically distributed normal as $N(\phi_{ij}, \sigma_{ij}^2)$ independently. Thus, the measure Φ has an asymptotically normal distribution with mean

$$\sum_{i < j} \sum w_{ij} \phi_{ij}, \tag{14}$$

and variance

$$\sigma^2 = \sum_{i < j} \sum (w_{ij})^2 \sigma_{ij}^2. \quad (15)$$

In an analogous manner to Agresti [12] [p. 170], we derive the weights $\{w_{ij}^*\}$ so as to minimize the variance of Φ with the constraint that all $w_{ij} > 0$ and $\sum_{i < j} \sum w_{ij} = 1$. From Appendix B, we obtain

$$w_{ij}^* = \frac{1/\sigma_{ij}^2}{\sum_{s < t} \sum 1/\sigma_{st}^2}. \quad (16)$$

Then, we consider the following measure, which represents the degree of departure from symmetry:

$$\Phi_S = \sum_{i < j} \sum w_{ij}^* \hat{\phi}_{ij}. \quad (17)$$

The measure (17) has the smallest variance among measures in the class of weighted averages given in Equation (13). It should be noted that we should estimate the variances $\{\sigma_{ij}^2\}$ because these are unknown.

We propose the estimated measure as follows:

$$\hat{\Phi}_S = \sum_{i < j} \sum \hat{w}_{ij}^* \hat{\phi}_{ij}, \quad (18)$$

where \hat{w}_{ij}^* is given by w_{ij}^* with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$. The proposed measure approximates the measure in the class of weighted averages that has the smallest variance. The estimated measure $\hat{\Phi}_T$ is the weighted average of $\hat{\phi}_{ij}$ using the weights $\{\hat{w}_{ij} = (\hat{p}_{ij} + \hat{p}_{ji})/\hat{\delta}\}$, where $\hat{\delta} = \sum_{i \neq j} \hat{p}_{ij}$. On the other hand, the proposed measure $\hat{\Phi}_S$ is the weighted average of $\hat{\phi}_{ij}$ using the weights $\{\hat{w}_{ij}^*\}$. It should be noted that (i) $\hat{w}_{ij} = (\hat{p}_{ij} + \hat{p}_{ji})/\hat{\delta}$ indicates the estimated conditional probability that the observation falls in (i, j) or (j, i) cells on the condition that the observation falls in off-diagonal cells and (ii) the weight \hat{w}_{ij}^* becomes larger as the variance of partial measure $\hat{\phi}_{ij}$ decreases.

4. Numerical Examples

The objective is to confirm the difference in the single summary measure for symmetry by comparing the weights $\{\hat{w}_{ij}\}$ and $\{\hat{w}_{ij}^*\}$. Consider the artificial data in Table 1(a)–(d) with $n = 1000$ and Table 1(e) with $n = 200$. Table 1(a)–(d) are generated from the random numbers of the multinomial distribution based on the cell probability tables (a), (b), (c), and (d) in Table 2, respectively. Table 1(e) is generated from the random numbers of the multinomial distribution based on the cell probability table in Table 2(a). The artificial cell probability tables of Table 2 focus in particular on the probabilities of cells (1, 2) and (2, 1), and the four patterns (a), (b), (c), and (d) are set according to the combination of partial symmetry/asymmetry. We shall apply the partial measure ϕ_{ij} . Table 3 shows the estimated partial measure $\hat{\phi}_{ij}$, estimated variance $\hat{\sigma}_{ij}^2$, confidence interval for ϕ_{ij} , estimated weights $\{\hat{w}_{ij}\}$ for $\hat{\Phi}_T$, and estimated weights $\{\hat{w}_{ij}^*\}$ for $\hat{\Phi}_S$. Figure 1 visualizes the estimated partial measure $\hat{\phi}_{ij}$ and confidence interval for ϕ_{ij} . The confidence interval for ϕ_{12} applied to the data in Table 1(a) does not contain zero, indicating that there is a partially asymmetric structure in cells (1, 2) and (2, 1). Furthermore, the confidence interval for ϕ_{12} does not overlap with the confidence intervals for ϕ_{ij} for any other pair of cells, indicating that cells (1, 2) and (2, 1) are partially asymmetric compared to every other pair of cells. The \hat{w}_{12}^* of $\hat{\Phi}_S$ is remarkably smaller than the \hat{w}_{12} of $\hat{\Phi}_T$, and the value of $\hat{\Phi}_S$ is smaller than that of

$\hat{\Phi}_T$. The value of $\hat{\phi}_{12}$ applied to the data in Table 1(b) is as large as the value of $\hat{\phi}_{12}$ applied to the data in Table 1(a). However, it cannot be shown that there is a partially asymmetric structure in cells (1,2) and (2,1) in Table 1(b) because the confidence interval for ϕ_{12} is wide and contains zero. Both $\hat{\Phi}_S$ and $\hat{\Phi}_T$ in Table 1(b) have small values of \hat{w}_{12} and \hat{w}_{12}^* , and in particular, the \hat{w}_{12}^* for $\hat{\Phi}_S$ is remarkably small.

Table 1. Artificial data.

(a)	(1)	(2)	(3)	(4)	(5)
(1)	37	544	12	7	8
(2)	102	26	15	15	12
(3)	9	8	29	10	11
(4)	9	9	12	40	12
(5)	14	9	10	11	29
(b)	(1)	(2)	(3)	(4)	(5)
(1)	47	11	37	44	48
(2)	3	38	34	37	49
(3)	44	44	52	56	48
(4)	38	25	55	45	47
(5)	35	25	51	43	44
(c)	(1)	(2)	(3)	(4)	(5)
(1)	33	316	13	18	18
(2)	321	37	20	18	20
(3)	7	6	26	16	14
(4)	5	5	1	30	19
(5)	5	10	5	2	35
(d)	(1)	(2)	(3)	(4)	(5)
(1)	39	4	70	42	50
(2)	5	34	45	110	84
(3)	17	12	54	103	63
(4)	31	14	20	39	48
(5)	9	29	26	6	46
(e)	(1)	(2)	(3)	(4)	(5)
(1)	7	103	1	1	4
(2)	19	10	2	2	4
(3)	2	1	6	2	2
(4)	3	4	4	5	2
(5)	1	2	4	1	8

Table 2. Artificial cell probability tables.

(a)	(1)	(2)	(3)	(4)	(5)
(1)	0.030	0.570	0.010	0.010	0.010
(2)	0.010	0.030	0.010	0.010	0.010
(3)	0.010	0.010	0.030	0.010	0.010
(4)	0.010	0.010	0.010	0.030	0.010
(5)	0.010	0.010	0.010	0.010	0.030

Table 2. Cont.

	(b)	(1)	(2)	(3)	(4)	(5)
(1)		0.040	0.008	0.040	0.040	0.040
(2)		0.004	0.040	0.040	0.040	0.040
(3)		0.040	0.040	0.050	0.050	0.050
(4)		0.040	0.040	0.050	0.040	0.049
(5)		0.040	0.040	0.050	0.049	0.040
	(c)	(1)	(2)	(3)	(4)	(5)
(1)		0.030	0.320	0.015	0.015	0.015
(2)		0.320	0.030	0.021	0.019	0.024
(3)		0.005	0.007	0.030	0.020	0.016
(4)		0.005	0.006	0.005	0.030	0.018
(5)		0.005	0.007	0.004	0.003	0.030
	(d)	(1)	(2)	(3)	(4)	(5)
(1)		0.040	0.003	0.080	0.050	0.050
(2)		0.003	0.040	0.050	0.100	0.080
(3)		0.020	0.010	0.050	0.100	0.064
(4)		0.030	0.020	0.020	0.040	0.050
(5)		0.010	0.020	0.020	0.020	0.040

Table 3. Estimate of measure ϕ_{ij} , estimated approximate variance for ϕ_{ij} , approximate 95% confidence interval for ϕ_{ij} , weights for measures Φ_S and Φ_T , and estimates of measures Φ_S and Φ_T , applied to Table 1(a)–(d).

Applied Data	Cells	$\hat{\phi}_{ij}$	$\hat{\sigma}_{ij}^2$	Confidence Interval for ϕ_{ij}	\hat{w}_{ij}^*	\hat{w}_{ij}	$\hat{\Phi}_S$	$\hat{\Phi}_T$
Table 1(a)	(1,2), (2,1)	0.37075	0.0012005	(0.303, 0.439)	0.057987	0.769964	0.02624	0.29124
	(1,3), (3,1)	0.01477	0.0020088	(−0.073, 0.103)	0.034653	0.025030		
	(1,4), (4,1)	0.01130	0.0020219	(−0.077, 0.099)	0.034428	0.019070		
	(1,5), (5,1)	0.05434	0.0068561	(−0.108, 0.217)	0.010153	0.026222		
	(2,3), (3,2)	0.06789	0.0081116	(−0.109, 0.244)	0.008582	0.027414		
	(2,4), (4,2)	0.04557	0.0053039	(−0.097, 0.188)	0.013124	0.028605		
	(2,5), (5,2)	0.01477	0.0020088	(−0.073, 0.103)	0.034653	0.025030		
	(3,4), (4,3)	0.00597	0.0007797	(−0.049, 0.061)	0.089277	0.026222		
	(3,5), (5,3)	0.00164	0.0002246	(−0.028, 0.031)	0.309966	0.025030		
	(4,5), (5,4)	0.00136	0.0001710	(−0.024, 0.027)	0.407178	0.027414		
Table 1(b)	(1,2), (2,1)	0.25041	0.0422558	(−0.152, 0.653)	0.000032	0.018088	0.00036	0.01843
	(1,3), (3,1)	0.00539	0.0001914	(−0.022, 0.033)	0.006979	0.104651		
	(1,4), (4,1)	0.00387	0.0001357	(−0.019, 0.027)	0.009848	0.105943		
	(1,5), (5,1)	0.01777	0.0006101	(−0.031, 0.066)	0.002190	0.107235		
	(2,3), (3,2)	0.01189	0.0004362	(−0.029, 0.053)	0.003063	0.100775		
	(2,4), (4,2)	0.02719	0.0012416	(−0.042, 0.096)	0.001076	0.080103		
	(2,5), (5,2)	0.07727	0.0028494	(−0.027, 0.182)	0.000469	0.095607		
	(3,4), (4,3)	0.00006	0.0000015	(−0.002, 0.002)	0.877858	0.143411		
	(3,5), (5,3)	0.00066	0.0000193	(−0.008, 0.009)	0.069221	0.127907		
	(4,5), (5,4)	0.00143	0.0000457	(−0.012, 0.015)	0.029264	0.116279		

Table 3. Cont.

Applied Data	Cells	$\hat{\phi}_{ij}$	$\hat{\sigma}_{ij}^2$	Confidence Interval for ϕ_{ij}	\hat{w}_{ij}^*	\hat{w}_{ij}	$\hat{\Phi}_S$	$\hat{\Phi}_T$
Table 1(c)	(1,2), (2,1)	0.00004	0.0000002	(−0.001, 0.001)	0.999900	0.759237	0.00006	0.06270
	(1,3), (3,1)	0.06593	0.0090727	(−0.121, 0.253)	0.000022	0.023838		
	(1,4), (4,1)	0.24463	0.0252616	(−0.067, 0.556)	0.000008	0.027414		
	(1,5), (5,1)	0.24463	0.0252616	(−0.067, 0.556)	0.000008	0.027414		
	(2,3), (3,2)	0.22065	0.0205989	(−0.061, 0.502)	0.000010	0.030989		
	(2,4), (4,2)	0.24463	0.0252616	(−0.067, 0.556)	0.000008	0.027414		
	(2,5), (5,2)	0.08170	0.0074074	(−0.087, 0.250)	0.000027	0.035757		
	(3,4), (4,3)	0.67724	0.0521067	(0.230, 1.125)	0.000004	0.020262		
	(3,5), (5,3)	0.16853	0.0225185	(−0.126, 0.463)	0.000009	0.022646		
(4,5), (5,4)	0.54628	0.0432851	(0.139, 0.954)	0.000005	0.025030			
Table 1(d)	(1,2), (2,1)	0.00892	0.0028433	(−0.096, 0.113)	0.113903	0.011421	0.11518	0.28830
	(1,3), (3,1)	0.28736	0.0075340	(0.117, 0.457)	0.042986	0.110406		
	(1,4), (4,1)	0.01644	0.0006424	(−0.033, 0.066)	0.504107	0.092640		
	(1,5), (5,1)	0.38383	0.0134101	(0.157, 0.611)	0.024150	0.074873		
	(2,3), (3,2)	0.25751	0.0106028	(0.056, 0.459)	0.030545	0.072335		
	(2,4), (4,2)	0.49139	0.0071440	(0.326, 0.657)	0.045333	0.157360		
	(2,5), (5,2)	0.17837	0.0039745	(0.055, 0.302)	0.081484	0.143401		
	(3,4), (4,3)	0.35950	0.0061895	(0.205, 0.514)	0.052323	0.156091		
	(3,5), (5,3)	0.12854	0.0037881	(0.008, 0.249)	0.085494	0.112944		
(4,5), (5,4)	0.49674	0.0164609	(0.245, 0.748)	0.019674	0.068528			
Table 1(e)	(1,2), (2,1)	0.37599	0.0064089	(0.219, 0.533)	0.486330	0.743902	0.23244	0.30922
	(1,3), (3,1)	0.08170	0.0740741	(−0.452, 0.615)	0.042077	0.018293		
	(1,4), (4,1)	0.18872	0.1177550	(−0.484, 0.861)	0.026469	0.024390		
	(1,5), (5,1)	0.27807	0.1280000	(−0.423, 0.979)	0.024350	0.030488		
	(2,3), (3,2)	0.08170	0.0740741	(−0.452, 0.615)	0.042077	0.018293		
	(2,4), (4,2)	0.08170	0.0370370	(−0.295, 0.459)	0.084155	0.036585		
	(2,5), (5,2)	0.08170	0.0370370	(−0.295, 0.459)	0.084155	0.036585		
	(3,4), (4,3)	0.08170	0.0370370	(−0.295, 0.459)	0.084155	0.036585		
	(3,5), (5,3)	0.08170	0.0370370	(−0.295, 0.459)	0.084155	0.036585		
(4,5), (5,4)	0.08170	0.0740741	(−0.452, 0.615)	0.042077	0.018293			

On the other hand, the value of $\hat{\phi}_{12}$ applied to the data in Table 1(c) indicates that there is a partially symmetric structure in cells (1, 2) and (2, 1) because the value of $\hat{\phi}_{12}$ is small and the confidence interval for the ϕ_{12} contains zero. In addition, the \hat{w}_{12}^* in Table 1(c) is large, indicating that the weight of $\hat{\Phi}_S$ is larger when the pair of cells is more frequent than others and has a partially symmetric structure. Both $\hat{\Phi}_S$ and $\hat{\Phi}_T$ applied to the data in Table 1(c) are close to zero because of the greater weight of the pair of cells (1, 2) and (2, 1) that show partial symmetry compared to the pairs of cells (4, 5) and (5, 4) and (3, 4) and (4, 3) that show partial asymmetry. The values of $\hat{\phi}_{ij}$ applied to the data in Table 1(d) indicate that the pair of cells (1, 2) and (2, 1) and the pair of cells (1, 4) and (4, 1) both have a partially symmetric structure because the confidence intervals of ϕ_{12} and ϕ_{14} include zero. It can be seen that the values of \hat{w}_{12} and \hat{w}_{14} for $\hat{\Phi}_T$ are similar, while the value of \hat{w}_{14}^* for $\hat{\Phi}_S$ is large compared to \hat{w}_{12}^* .

The value of $\hat{\phi}_{12}$ applied to the data in Table 1(e) is about the same as the value applied to the data in Table 1(a), but the confidence interval is wider due to the smaller sample size, making it relatively difficult to conclude partial asymmetry for the pair of cells (1, 2) and (2, 1). The magnitude of $\hat{\phi}_{12}$ applied to Table 1(e) does not differ much from the results applied to Table 1(a). However, the values of \hat{w}_{12}^* and $\hat{\Phi}_S$ are greater when applied to Table 1(e) than Table 1(a). It should be noted that the weight \hat{w}_{ij}^* becomes larger as the variance of the partial measure $\hat{\phi}_{ij}$ decreases, and the weight \hat{w}_{ij} becomes larger as the proportion $(n_{ij} + n_{ji})/n$ increases.

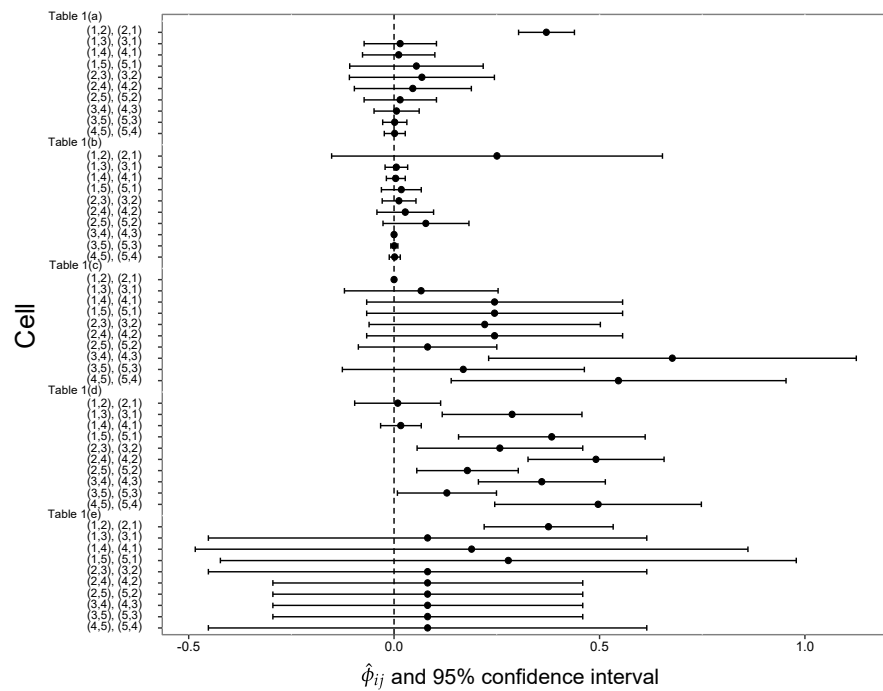


Figure 1. Estimate of measure ϕ_{ij} and approximate 95% confidence interval for ϕ_{ij} applied to Table 1.

5. Example

Consider the data in Table 4, derived from the national survey on educational attitudes of high school students and their mothers in Japan in 2012. To clarify the structure of educational inequalities in contemporary Japanese society and the actual educational awareness of parents and children, a postal survey was conducted among second-year high school students and their mothers throughout Japan, using the same framework as the national survey on the educational awareness of high school students and their mothers conducted in November 2002. The data describe the cross-classification of mothers’ and fathers’ birth orders. For example, for the 179 high school students whose mothers’ birth order is “First” and whose fathers’ birth order is “Second”, the mother is the eldest daughter and the father is the second son. The partial symmetry of cells (1,2) and (2,1) means that the probability of high school students whose mother is the eldest daughter and whose father is the second son is equal to that of high school students whose mother is the second daughter and whose father is the eldest son.

Table 4. Cross-classification of mothers’ and fathers’ birth orders.

Mothers’ Birth Order	Fathers’ Birth Order				Total
	First	Second	Third	Fourth or More	
First	224	179	53	22	478
Second	162	153	35	15	365
Third	37	37	18	11	103
Fourth or more	12	7	3	5	27
Total	435	376	109	53	973

Let G^2_ζ and X^2_ζ denote the likelihood ratio and Pearson’s chi-squared statistics for testing the goodness of fit of the symmetry model, i.e., $G^2_\zeta = 2 \sum \sum_{i \neq j} n_{ij} \log(2n_{ij}/(n_{ij} + n_{ji}))$ and $X^2_\zeta = \sum \sum_{i < j} (n_{ij} - n_{ji})^2 / (n_{ij} + n_{ji})$. For large samples, G^2_ζ and X^2_ζ have a chi-squared null distribution with $r(r - 1)/2$ degrees of freedom. From $G^2_\zeta = 14.58$ and $X^2_\zeta = 14.17$ with six degrees of freedom for the data in Table 4, these values indicate the lack of a symmetrical structure. Note that the exact test introduced by West [13] is well known as a

test for the contingency table including structural zeros. As the proposed measure does not require the frequency of the diagonal components, West's test was also conducted assuming that the diagonal components are structural zeros. The simulated p-value from West's test is 0.085, which indicates that the rows and columns are independent. The value of $\hat{\Phi}_T$ is 0.0184, and the 95% confidence interval is (0.000003, 0.036719), which does not include zero.

Next, we measured the degree of departure from partial symmetry for each pair of cells. We shall apply the partial measure ϕ_{ij} for the data in Table 4. Table 5 shows the estimated values for ϕ_{ij} and σ_{ij}^2 , confidence intervals for ϕ_{ij} , estimated weights $\{\hat{w}_{ij}\}$ and $\{\hat{w}_{ij}^*\}$, and estimated measures $\hat{\Phi}_S$ and $\hat{\Phi}_T$. According to the magnitudes of the estimates, ϕ_{ij} can explain the partial symmetry for each pair of cells in Table 4. The 95% confidence interval for the ϕ_{ij} for all pairs of cells contains zero, which indicates that there is a partially symmetrical structure in each birth order category in the mother–father pairs.

Table 5. Estimate of measure ϕ_{ij} , estimated approximate variance for ϕ_{ij} , approximate 95% confidence interval for ϕ_{ij} , estimates of measures of Φ_S and Φ_T , and weights for measures of Φ_S and Φ_T , applied to Table 4.

Cells	$\hat{\phi}_{ij}$	$\hat{\sigma}_{ij}^2$	Confidence Interval for ϕ_{ij}	\hat{w}_{ij}^*	\hat{w}_{ij}	$\hat{\Phi}_S$	$\hat{\Phi}_T$
(1,2), (2,1)	0.0018	0.00002	(−0.006, 0.009)	0.5864	0.5951		
(1,3), (3,1)	0.0229	0.00072	(−0.030, 0.076)	0.0123	0.1571		
(1,4), (4,1)	0.0633	0.00514	(−0.077, 0.204)	0.0017	0.0593		
(2,3), (3,2)	0.0006	0.00002	(−0.009, 0.010)	0.3986	0.1257	0.0018	0.0184
(2,4), (4,2)	0.0976	0.01192	(−0.116, 0.312)	0.0007	0.0384		
(3,4), (4,3)	0.2504	0.04226	(−0.152, 0.653)	0.0002	0.0244		

Furthermore, the estimated departure from symmetry is smaller with $\hat{\Phi}_S$, which uses different weights than $\hat{\Phi}_T$. Figure 2 plots estimated weights \hat{w}_{ij} and \hat{w}_{ij}^* . Cells (1,2) and (2,1) have similar frequencies and are more frequent than the other cells. Then, weights \hat{w}_{ij} and \hat{w}_{ij}^* are similar and large. On the other hand, the pair of cells (2,3) and (3,2) have similar frequencies, but are less frequent than the pair of cells (1,2) and (2,1). In such cases, \hat{w}_{ij}^* is larger than \hat{w}_{ij} . Therefore, $\hat{\Phi}_S$ has a higher weight than $\hat{\Phi}_T$ when the pair of cells has a lower frequency than another pair of cells and when the cells have similar frequencies. Conversely, \hat{w}_{ij}^* is smaller than \hat{w}_{ij} when the frequencies are different, as in the pair of cells (1,3) and (3,1). Since the weights $\{\hat{w}_{ij}^*\}$ and $\{\hat{w}_{ij}\}$ take different values, the single summary measures $\hat{\Phi}_S$ and $\hat{\Phi}_T$ also take different values. As mentioned above, there is a partially symmetrical structure in each birth order category in the mother–father pairs. Then, the proposed measure may be reasonable to express the degree of departure from symmetry.

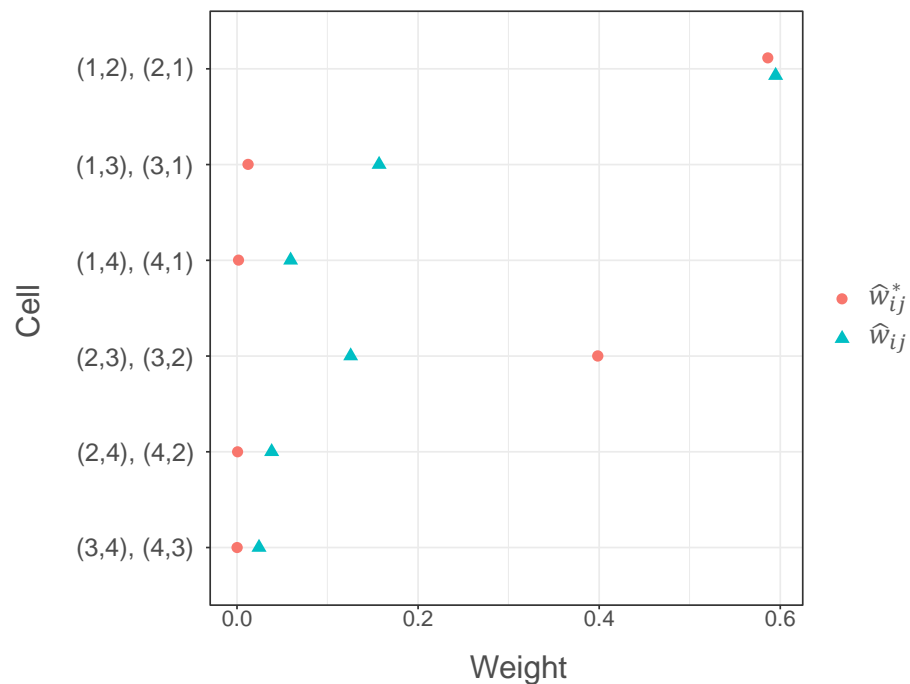


Figure 2. The weight for each pair of symmetric cells obtained by applying the proposed method and Tomizawa [10] to Table 4.

6. Concluding Remarks

We proposed a partial measure to express the degree of departure from partial symmetry. The measure was constructed as the weighted average of partial measures expressed using the Shannon entropy or Kullback–Leibler information. The composition of the proposed measure Φ_S is similar to that of the measure proposed by Tomizawa [10] in the sense that they are classes of weighted averages. However, they differ in that the weights multiplied by the partial measure are constructed so as to minimize the measure's variance. This measure increase with the degree of departure from symmetry, allowing us to see how far away the probability structure of the contingency table is from complete asymmetry.

The measures $\hat{\Phi}_S$ and $\hat{\Phi}_T$ are invariant under the arbitrary simultaneous permutations of row and column categories, and therefore, it is possible to apply these measures to analyze the data on a nominal scale, as well as on an ordinal scale if one cannot use the information about the order in which the categories are listed.

We compared the weights used to construct the measures $\hat{\Phi}_S$ and $\hat{\Phi}_T$. Those used to construct $\hat{\Phi}_T$ are large when the frequency of the pair of cells is high compared to others. On the other hand, the weight of $\hat{\Phi}_S$ is higher when the frequency of the pair of cells is higher than others and when the structure is partially symmetric. Conversely, when the frequency of the pair of cells is lower than others and the structure is partially asymmetric, the weights of $\hat{\Phi}_S$ are smaller than those of $\hat{\Phi}_T$.

In the present study, confidence intervals for partial measures were used to interpret the partially asymmetric structure of the data. Alternatively, global tests for the null hypothesis that all ϕ_{ij} are equally zero, and multiplicity correction for paired comparisons also need to be considered and are left as future works.

We should note, however, that $\hat{\Phi}_S$ cannot be calculated if any of the off-diagonal cells are zero. As such, the proposed measure should be used for contingency tables with large sample sizes.

Author Contributions: Conceptualization, T.I., K.T. and K.Y.; methodology, T.I., K.T. and K.Y.; software, T.I.; validation, K.T.; formal analysis, T.I.; investigation, T.I.; writing—original draft preparation, T.I., K.T. and K.Y.; writing—review and editing, T.I., K.T., K.Y. and S.T.; visualization, T.I.; supervision, S.T.; project administration, K.T.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: The data, for example, in Section 5, the national survey on educational attitudes of high school students and their mothers in Japan in 2012, were provided by the Social Science Japan Data Archive, Center for Social Research and Data Archives, Institute of Social Science, The University of Tokyo. Furthermore, the authors are grateful to an Associate Editor and three Referees for their constructive comments and useful suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

From the central limit theorem, $\hat{\boldsymbol{p}}$ is asymptotically distributed as normal $N(\boldsymbol{p}, \Sigma_1(\boldsymbol{p}))$, where $\Sigma_1(\boldsymbol{p})$ is the $r^2 \times r^2$ matrix

$$\Sigma_1(\boldsymbol{p}) = \frac{1}{n}(D(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}'), \quad (\text{A1})$$

where $D(\boldsymbol{p})$ denotes a diagonal matrix with the i th element of \boldsymbol{p} as the i th diagonal element. Then, we also obtain

$$\hat{\boldsymbol{\phi}} = \boldsymbol{\phi} + d_1(\boldsymbol{p})(\hat{\boldsymbol{p}} - \boldsymbol{p}) + o(\|\hat{\boldsymbol{p}} - \boldsymbol{p}\|), \quad (\text{A2})$$

where $d_1(\boldsymbol{p}) = \partial\boldsymbol{\phi}/\partial\boldsymbol{p}'$ is the $r(r-1)/2 \times r^2$ matrix. Thus, $\hat{\boldsymbol{\phi}}$ is asymptotically distributed as normal $N(\boldsymbol{\phi}, \sigma^2[\hat{\boldsymbol{\phi}}])$, where

$$\sigma^2[\hat{\boldsymbol{\phi}}] = d_1(\boldsymbol{p})\Sigma_1(\boldsymbol{p})d_1(\boldsymbol{p})'. \quad (\text{A3})$$

Let $\hat{\boldsymbol{\pi}}'$ be the $1 \times r(r-1)$ vector:

$$\hat{\boldsymbol{\pi}}' = (\hat{\pi}'_{(12)}, \hat{\pi}'_{(13)}, \dots, \hat{\pi}'_{(r-1,r)}), \quad (\text{A4})$$

where $\hat{\pi}'_{(ij)} = (\hat{\pi}_{ij}, \hat{\pi}_{ji})$. Noting that $\hat{\boldsymbol{\phi}}$ is a function of only $\{\pi_{ij}\}$, we obtain

$$d_1(\boldsymbol{p}) = \frac{\partial\boldsymbol{\phi}}{\partial\boldsymbol{\pi}'} \cdot \frac{\partial\boldsymbol{\pi}}{\partial\boldsymbol{p}'}. \quad (\text{A5})$$

It should be noted that $\partial\boldsymbol{\phi}/\partial\boldsymbol{\pi}'$ is the $r(r-1)/2 \times r(r-1)$ matrix and $\partial\boldsymbol{\pi}/\partial\boldsymbol{p}'$ is the $r(r-1) \times r^2$ matrix. By obtaining $\partial\boldsymbol{\pi}/\partial\boldsymbol{p}'$, we can see that $\sigma^2[\hat{\boldsymbol{\phi}}]$ is expressed as

$$\sigma^2[\hat{\boldsymbol{\phi}}] = \left(\frac{\partial\boldsymbol{\phi}}{\partial\boldsymbol{\pi}'}\right) \cdot \Sigma_2(\boldsymbol{p}) \cdot \left(\frac{\partial\boldsymbol{\phi}}{\partial\boldsymbol{\pi}'}\right)', \quad (\text{A6})$$

where $\Sigma_2(\boldsymbol{p})$ is the $r(r-1) \times r(r-1)$ matrix:

$$\Sigma_2(\boldsymbol{p}) = \begin{pmatrix} \Sigma_{12}(\boldsymbol{p}) & 0 & \cdots & 0 \\ 0 & \Sigma_{13}(\boldsymbol{p}) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \Sigma_{r-1,r}(\boldsymbol{p}) \end{pmatrix}, \quad (\text{A7})$$

where

$$\Sigma_{ij}(\mathbf{p}) = \frac{1}{n(p_{ij} + p_{ji})} \begin{pmatrix} \pi_{ij}(1 - \pi_{ij}) & -\pi_{ij}\pi_{ji} \\ -\pi_{ij}\pi_{ji} & \pi_{ji}(1 - \pi_{ji}) \end{pmatrix} \quad (i < j). \quad (\text{A8})$$

Thus, $\sigma^2[\hat{\boldsymbol{\phi}}]$ is also expressed as follows:

$$\sigma^2[\hat{\boldsymbol{\phi}}] = \begin{pmatrix} \sigma_{12}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{13}^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{r-1,r}^2 \end{pmatrix}, \quad (\text{A9})$$

where

$$\sigma_{ij}^2 = \frac{\pi_{ij}\pi_{ji}}{n(p_{ij} + p_{ji})} \left(\frac{1}{\log 2} (\log \pi_{ij} - \log \pi_{ji}) \right)^2 \quad (i < j). \quad (\text{A10})$$

Appendix B

Let \mathbf{w} be the $r(r-1)/2 \times 1$ vector:

$$\mathbf{w} = (w_{12}, w_{13}, \dots, w_{r-1,r})'. \quad (\text{A11})$$

Then, the measure (13) is expressed as $\Phi = \mathbf{w}'\hat{\boldsymbol{\phi}}$.

From Appendix A, the mean and variance of Φ are approximately calculated as follows:

$$E(\Phi) = \mathbf{w}'\boldsymbol{\phi}, \quad (\text{A12})$$

$$\text{Var}(\Phi) = \mathbf{w}'\sigma^2[\hat{\boldsymbol{\phi}}]\mathbf{w}. \quad (\text{A13})$$

Then, we can obtain the following \mathbf{w}^* so as to minimize $\text{Var}(\Phi)$ with the constraint that $\mathbf{w}'\mathbf{1}_d$ ($d = r(r-1)/2$) is unity ($\mathbf{1}_d$ is the $d \times 1$ vector of 1 elements):

$$w_{ij}^* = \frac{1/\sigma_{ij}^2}{\sum_{s < t} 1/\sigma_{st}^2} \quad (i < j). \quad (\text{A14})$$

References

1. Agresti, A. *Categorical Data Analysis*, 3rd ed.; Wiley: Hoboken, NJ, USA, 2013.
2. Bishop, Y.M.M.; Fienberg, S.E.; Holl, P.W. *Discrete Multivariate Analysis: Theory and Practice*; The MIT Press: Cambridge, MA, USA, 1975.
3. Beh, E.J.; Simonetti, B.; D'Ambra, L. Partitioning a non-symmetric measure of association for three-way contingency tables. *J. Multivar. Anal.* **2007**, *98*, 1391–1411. [[CrossRef](#)]
4. Lombardo, R. Three-way association measure decompositions: The Delta index. *J. Stat. Plan. Inference* **2011**, *141*, 1789–1799. [[CrossRef](#)]
5. Wei, Z.; Kim, D. Subcopula-based measure of asymmetric association for contingency tables. *Stat. Med.* **2017**, *36*, 3875–3894. [[CrossRef](#)] [[PubMed](#)]
6. Wei, Z.; Kim, D. Measure of asymmetric association for ordinal contingency tables via the bilinear extension copula. *Stat. Probab. Lett.* **2021**, *178*, 109183. [[CrossRef](#)]
7. Zhang, L.; Lu, D.; Wang, X. The essential dependence for a group of random vectors. *Commun. Stat.-Theory Methods* **2021**, *50*, 5836–5872. [[CrossRef](#)]
8. Wei, Z.; Kim, D.; Conlon, E.M. A Bayesian approach to the analysis of asymmetric association for two-way contingency tables. *Comput. Stat.* **2022**, *37*, 1311–1338. [[CrossRef](#)]
9. Bowker, A.H. A test for symmetry in contingency tables. *J. Am. Stat. Assoc.* **1948**, *43*, 572–574. [[CrossRef](#)] [[PubMed](#)]
10. Tomizawa, S. Two kinds of measures of departure from symmetry in square contingency tables having nominal categories. *Stat. Sin.* **1994**, *4*, 325–334.
11. Fernandes, L.H.S.; Araújo, F.H.A. Taxonomy of commodities assets via the complexity-entropy causality plane. *Chaos Solitons Fractals* **2020**, *137*, 109909. [[CrossRef](#)]

12. Agresti, A. *Analysis of Ordinal Categorical Data*; Wiley: New York, NY, USA, 1984.
13. West, L.J.; Hankin, R.K.S. Exact tests for two-way contingency tables with structural zeros. *J. Stat. Softw.* **2008**, *28*, 1–19. [[CrossRef](#)]