symmetry

MDPI

Article

Measure of Departure from Conditional Symmetry Based on Cumulative Probabilities for Square Contingency Tables

Yusuke Saigusa [1,*], Yuta Teramoto [2] and Sadao Tomizawa [2]

[1] Department of Biostatistics, Yokohama City University Graduate School of Medicine, Yokohama 236-0004, Kanagawa, Japan
[2] Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda 278-8510, Chiba, Japan; 6316626@ed.tus.ac.jp (Y.T.); tomizawa@is.noda.tus.ac.jp (S.T.)
* Correspondence: saigusay@yokohama-cu.ac.jp

Abstract: For the analysis of square contingency tables with ordered categories, a measure was developed to represent the degree of departure from the conditional symmetry model in which there is an asymmetric structure of the cell probabilities with respect to the main diagonal of the table. The present paper proposes a novel measure for the departure from conditional symmetry based on the cumulative probabilities from the corners of the square table. In a given example, the proposed measure is applied to Japanese occupational status data, and the interpretation of the proposed measure is illustrated as the departure from a proportional structure of social mobility.

Keywords: conditional symmetry; contingency table; diversity index; measure; ordered category

Citation: Saigusa, Y.; Teramoto, Y.; Tomizawa, S. Measure of Departure from Conditional Symmetry Based on Cumulative Probabilities for Square Contingency Tables. *Symmetry* **2021**, *13*, 1897. https://doi.org/10.3390/sym13101897

Academic Editors: Jinyu Li and Dalibor Štys

Received: 17 August 2021
Accepted: 4 October 2021
Published: 8 October 2021

1. Introduction

Symmetry and asymmetry issues frequently arise in square contingency tables with the same row and column classifications from a broad range of scientific fields, for example, medical, social, and geographical sciences [1,2]. A number of statistical approaches were developed to find the symmetric and asymmetric relationships of the underlying probability distribution for the square table. The authors in [3] considered the symmetry model for cell (joint) probability in the square table, and, following that, various symmetry and asymmetry models were investigated and evaluated through data analysis [4,5]. The researchers in [1,6,7] considered the singular value decomposition of a skew-symmetric matrix, which is composed of the residuals from the symmetry model. Regarding the residuals from the symmetry model, [8] gave the index, which represents the degree of residuals, and [9] considered the correspondence analysis of the residual matrix. The present paper focuses on measuring the degree of departure from an asymmetric structure for cumulative cell probabilities.

Consider an $R \times R$ contingency table with the same row and column classifications with ordered categories. Let X and Y denote the row and column variables, respectively. Let $p_{ij} = \Pr(X = i, Y = j)$ for $i = 1, \ldots, R; j = 1, \ldots, R$. The authors in [10] considered the conditional symmetry (CS) model defined by

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

In particular, if $\Delta = 1$ holds in the CS model, the symmetry model ([3,4], p. 282) holds. The CS model indicates a structure in which a cell probability is proportional to the symmetric cell probability with respect to the main diagonal with the common ratio to all pairs of probabilities.

When the CS model does not hold, we are interested in measuring how far the probability distribution is distant from the CS model. The authors in [11] gave a measure to represent the degree of departure from the CS model. Let $\Phi^{(\lambda)}$ denote the measure (see

Appendix A for the details). The value of measure $\Phi^{(\lambda)}$ is not invariant for re-ordering the categories of table. Therefore, $\Phi^{(\lambda)}$ is appropriate for applying to a square table with ordered categories.

We introduce the cumulative probability from the upper-right and lower-left corners of the square table. Let

$$G_{ij} = \Pr(X \leq i, Y \geq j) = \sum_{s=1}^{i} \sum_{t=j}^{R} p_{st} \quad (i < j),$$

$$G_{ji} = \Pr(X \geq j, Y \leq i) = \sum_{s=j}^{R} \sum_{t=1}^{i} p_{st} \quad (i < j).$$

The restriction of the CS model can be expressed by using the cumulative probabilities, that is,

$$G_{ij} = \Delta G_{ji} \quad (i < j). \tag{2}$$

Thus, the CS model indicates that each cumulative probability is proportionate to the symmetric cumulative probability with the ratio of Δ.

The cross-classification of the father's and his son's occupational statuses in Japan was explored in 1955 and 1965 [12] (p. 151). The occupational status was classified into five categories: (1) capitalist; (2) new-middle; (3) working; (4) self-employed; and (5) farming. These categories were treated as an ordinal scale by some statisticians; see for example, [13].

Thus, these data are given in the form of 5×5 ordinal contingency tables. Noting that the row variable is the father's status and the column variable is the son's status, the father–son pairs in the off-diagonal cells indicate the social mobility to a different status between the father and his son.

For the analysis of these data, if there is a conditional symmetric structure of the cumulative probabilities underlying the occupational status data, there is a proportional relation in which the probability of social mobility from father's status i or below to son's status j or above is Δ times higher than that from father's status j or above to son's status i or below for any value of the parameter of Δ, $i < j; i = 1, \ldots, 4; j = 2, \ldots, 5$.

The present paper focuses on measuring the degree of departure from the proportional structure of the cumulative probabilities (for the example described above) and comparing the degrees of departure among different tables. The measure $\Phi^{(\lambda)}$ would be useful to determine the degree of departure from a conditional symmetric structure (1) for the cell probabilities. On the other hand, because the CS model can also be expressed as (2), we might also be interested in measuring the degree to which the cumulative probabilities $\{G_{ij}\}$ are distant from those with a conditional symmetric structure.

Such a measure should be expressed as a function of the cumulative probabilities. The present paper considers a new measure to represent the degree of departure from CS based on the cumulative probabilities $\{G_{ij}\}$. Such a new measure may be useful when we determine the structure of the cumulative probabilities underlying the data rather than the structure of cell probabilities.

The rest of this paper is organized as follows. In Section 2, we propose a new measure that expresses the degree of departure from the CS model. In Section 3, we obtain the large-sample confidence interval for the proposed measure. In Section 4, we apply the measure to the actual data of the occupational status of father–son pairs in Japan and illustrate the interpretation of the measure. Section 5 provides our discussion.

2. Measure

For the cumulative probabilities $\{G_{ij}\}$ for an $R \times R$ table, the CS model can be expressed as

$$G_{ij}^{U} = G_{ji}^{L} \quad (i < j),$$

where

$$G_{ij}^{U} = \frac{G_{ij}}{\Delta_U}, \qquad G_{ji}^{L} = \frac{G_{ji}}{\Delta_L},$$

$$\Delta_U = \sum\sum_{s<t} G_{st}, \qquad \Delta_L = \sum\sum_{s>t} G_{st}.$$

Assume that $\Delta_U > 0$, $\Delta_L > 0$ and $\{G_{ij} + G_{ji} > 0\}$. Let $Q_{ij} = (G_{ij}^{U} + G_{ji}^{L})/2$ for $i < j$. Consider the measure defined by

$$\Psi^{(\lambda)} = \sum\sum_{i<j} Q_{ij}\left(1 - \frac{\lambda 2^{\lambda}}{2^{\lambda}-1} H_{ij}^{(\lambda)}\right) \qquad (\lambda > -1),$$

where

$$H_{ij}^{(\lambda)} = \frac{1}{\lambda}\left[1 - (G_{ij}^{c})^{\lambda+1} - (G_{ji}^{c})^{\lambda+1}\right],$$

$$G_{ij}^{c} = \frac{G_{ij}^{U}}{G_{ij}^{U} + G_{ji}^{L}}, \qquad G_{ji}^{c} = \frac{G_{ji}^{L}}{G_{ij}^{U} + G_{ji}^{L}}.$$

Note that $H_{ij}^{(\lambda)}$ is the Patil and Taillie's [14] diversity index of degree λ, which is a real value chosen by the user. Which value of λ to use will be discussed in Section 5. The component $H_{ij}^{(\lambda)}$ reflects the degree to which the weighted cumulative probabilities (G_{ij}^{c}, G_{ji}^{c}) are distant from uniform, $(1/2, 1/2)$, and these are rescaled in order to normalize the value of measure. The measure $\Psi^{(\lambda)}$ is formulated by combining the rescaled indices into the weighted mean with the weights of Q_{ij}, which is the relative magnitude of the cumulative probabilities (G_{ij}, G_{ji}). The measure evaluated at $\lambda = 0$ is taken to be the limit as $\lambda \to 0$, namely,

$$\Psi^{(0)} = \sum\sum_{i<j} Q_{ij}\left(1 - \frac{1}{\log 2} H_{ij}^{(0)}\right),$$

where

$$H_{ij}^{(0)} = -G_{ij}^{c} \log G_{ij}^{c} - G_{ji}^{c} \log G_{ji}^{c}.$$

Note that $H_{ij}^{(0)}$ is the Shannon entropy.

It is satisfied that (i) $0 \le \Psi^{(\lambda)} < 1$ and (ii) $\Psi^{(\lambda)}$ achieves the minimum value of 0 if and only if $G_{ij}^{U} = G_{ji}^{L}$ $(i < j)$, that is, there is a structure of the CS model in the table. We provide a numerical experiment to determine the change of value of $\Psi^{(\lambda)}$ for departures from the CS model in Section 5. It can be seen that the value of measure $\Psi^{(\lambda)}$ is not invariant for re-ordering the categories of table and, thus, incorporates the information of the order of categories.

3. Confidence Interval of Measure

Let n_{ij} denote the observed frequency in the ith row and jth column of the table $(i = 1, \ldots, R; j = 1, \ldots, R)$. The sample version of $\Psi^{(\lambda)}$, say $\hat{\Psi}^{(\lambda)}$, can be obtained from $\{\hat{p}_{ij}\}$, where $\hat{p}_{ij} = n_{ij}/n$ and $n = \sum\sum n_{ij}$. Assume that $\{n_{ij}\}$ result from full multinomial sampling. We shall consider an approximate standard error of $\hat{\Psi}^{(\lambda)}$ and a large-sample

confidence interval for $\Psi^{(\lambda)}$. Using the delta method [2] (p. 587), $\sqrt{n}(\hat{\Psi}^{(\lambda)} - \Psi^{(\lambda)})$ asymptotically (as $n \to \infty$) has a normal distribution with mean 0 and variance $\sigma^2\left[\Psi^{(\lambda)}\right]$ where

$$\sigma^2\left[\Psi^{(\lambda)}\right] = \sum_{i=1}^{R}\sum_{j=1}^{R} p_{ij}\left(\omega_{ij}^{(\lambda)}\right)^2 - \left(\sum_{i=1}^{R}\sum_{j=1}^{R} p_{ij}\omega_{ij}^{(\lambda)}\right)^2 \qquad (\lambda > -1),$$

with

$$\omega_{ij}^{(\lambda)} = \begin{cases} \sum\sum_{k<l}\left(\frac{2^\lambda(\lambda+1)}{2^\lambda-1}Q_{kl}\theta_{ijkl}\left((G_{kl}^c)^\lambda - (G_{lk}^c)^\lambda\right)\right. \\ \qquad \left. + \frac{\xi_{ijkl}+\xi_{ijlk}}{2}\left(1 - \frac{\lambda 2^\lambda}{2^\lambda-1}H_{kl}^{(\lambda)}\right)\right) & (\lambda \neq 0), \\ \sum\sum_{k<l}\left(\frac{1}{\log 2}Q_{kl}\theta_{ijkl}\log\left(\frac{G_{kl}^U}{G_{lk}^L}\right)\right. \\ \qquad \left. + \frac{\xi_{ijkl}+\xi_{ijlk}}{2}\left(1 - \frac{1}{\log 2}H_{kl}^{(0)}\right)\right) & (\lambda = 0), \end{cases}$$

$$\theta_{ijkl} = \frac{G_{lk}^c}{G_{kl}^U + G_{lk}^L}\xi_{ijkl} - \frac{G_{kl}^c}{G_{kl}^U + G_{lk}^L}\xi_{ijlk},$$

$$\xi_{ijkl} = \frac{I(i \le k, j \ge l)}{\Delta_U} - G_{kl}\frac{\sum\sum_{s<t} I(i \le s, j \ge t)}{\Delta_U^2},$$

$$\xi_{ijlk} = \frac{I(i \ge l, j \le k)}{\Delta_L} - G_{lk}\frac{\sum\sum_{s>t} I(i \ge s, j \le t)}{\Delta_L^2},$$

and $I(\cdot)$ is the indicator function. Let $\hat{\sigma}^2\left[\Psi^{(\lambda)}\right]$ denote $\sigma^2\left[\Psi^{(\lambda)}\right]$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$. We have the estimated approximate standard error of $\hat{\Psi}^{(\lambda)}$, $\hat{\sigma}\left[\Psi^{(\lambda)}\right]/\sqrt{n}$, and the approximate $100(1-p)\%$ confidence interval for $\Psi^{(\lambda)}$,

$$\left[\hat{\Psi}^{(\lambda)} - z_{p/2}\frac{\hat{\sigma}\left[\Psi^{(\lambda)}\right]}{\sqrt{n}}, \hat{\Psi}^{(\lambda)} + z_{p/2}\frac{\hat{\sigma}\left[\Psi^{(\lambda)}\right]}{\sqrt{n}}\right],$$

where $z_{p/2}$ is the $100(1-p/2)$th percentile of the standard normal distribution.

4. Data Analysis

Consider the Japanese father's and his son's occupational status data examined in 1955 and 1965 [12] (p. 151) introduced in Section 1 again. Applying the measure $\Psi^{(\lambda)}$ might be appropriate to evaluate the degree of departure from the proportional structure for the cumulative probability of social mobility and to compare the degrees of departure between in 1955 and in 1965. On the other hand, in order to evaluate the degree of departure from the proportional structure for the cell probabilities, the measure $\Phi^{(\lambda)}$ should be applied.

Table 1 gives the estimated values and the confidence intervals for $\Psi^{(\lambda)}$ applied to the occupational status data for some fixed λ. Similarly, Table 2 gives the estimated values and the confidence intervals for $\Phi^{(\lambda)}$. We compare the degrees of departure from CS between the occupational statuses in 1955 and in 1965 using the confidence intervals for $\Psi^{(\lambda)}$ given in Table 1. It is inferred that the degree of departure from CS in 1965 would be larger than that in 1955. Therefore, it can be interpreted from these results that the degree of departure from the proportional structure for the cumulative probabilities is larger for 1965 than for 1955. On the other hand, the results in Table 2 show that the degree of departure from the CS for the cell probabilities may be larger for 1965 than for 1955.

Table 1. Estimates of $\Psi^{(\lambda)}$, the estimated standard errors of $\hat{\Psi}^{(\lambda)}$, and 95% confidence intervals for $\Psi^{(\lambda)}$, applied to each of the occupational status data (**a**) in 1955 and (**b**) in 1965.

(**a**) Applied to the occupational status data in 1955.			
λ	**Estimate**	**Standard Error**	**Confidence Interval**
−0.8	0.023	0.004	(0.016, 0.030)
−0.6	0.042	0.006	(0.029, 0.055)
−0.4	0.058	0.009	(0.041, 0.075)
−0.2	0.071	0.010	(0.050, 0.091)
0	0.081	0.012	(0.058, 0.105)
0.2	0.090	0.014	(0.065, 0.115)
0.4	0.097	0.014	(0.070, 0.124)
0.6	0.102	0.014	(0.074, 0.130)
0.8	0.105	0.015	(0.076, 0.134)
1	0.108	0.015	(0.078, 0.138)
1.2	0.109	0.015	(0.080, 0.139)
1.4	0.110	0.015	(0.080, 0.140)
1.6	0.110	0.015	(0.080, 0.140)
(**b**) Applied to the occupational status data in 1965.			
λ	**Estimate**	**Standard Error**	**Confidence Interval**
−0.8	0.046	0.006	(0.033, 0.058)
−0.6	0.082	0.011	(0.061, 0.104)
−0.4	0.113	0.015	(0.084, 0.142)
−0.2	0.137	0.018	(0.103, 0.171)
0	0.156	0.020	(0.118, 0.195)
0.2	0.172	0.021	(0.130, 0.213)
0.4	0.183	0.022	(0.140, 0.227)
0.6	0.192	0.023	(0.147, 0.237)
0.8	0.198	0.023	(0.152, 0.244)
1	0.203	0.024	(0.156, 0.249)
1.2	0.205	0.024	(0.158, 0.252)
1.4	0.206	0.024	(0.159, 0.253)
1.6	0.206	0.024	(0.159, 0.253)

Table 2. Estimates of $\Phi^{(\lambda)}$, the estimated standard errors of $\hat{\Phi}^{(\lambda)}$, and 95% confidence intervals for $\Phi^{(\lambda)}$, applied to each of the occupational status data (**a**) in 1955 and (**b**) in 1965.

(**a**) Applied to the occupational status data in 1955.			
λ	**Estimate**	**Standard Error**	**Confidence Interval**
−0.8	0.058	0.008	(0.042, 0.074)
−0.6	0.104	0.014	(0.077, 0.131)
−0.4	0.140	0.018	(0.105, 0.175)
−0.2	0.168	0.021	(0.127, 0.209)
0	0.190	0.023	(0.145, 0.235)
0.2	0.207	0.025	(0.158, 0.255)
0.4	0.219	0.026	(0.169, 0.270)
0.6	0.229	0.026	(0.177, 0.281)
0.8	0.236	0.027	(0.183, 0.288)
1	0.240	0.027	(0.187, 0.294)
1.2	0.243	0.027	(0.189, 0.297)
1.4	0.244	0.028	(0.190, 0.298)
1.6	0.244	0.028	(0.190, 0.297)
(**b**) Applied to the occupational status data in 1965.			
λ	**Estimate**	**Standard Error**	**Confidence Interval**
−0.8	0.096	0.011	(0.074, 0.118)
−0.6	0.169	0.019	(0.132, 0.206)
−0.4	0.224	0.024	(0.178, 0.271)
−0.2	0.266	0.027	(0.214, 0.319)
0	0.298	0.029	(0.242, 0.355)
0.2	0.322	0.030	(0.263, 0.381)
0.4	0.340	0.031	(0.279, 0.401)
0.6	0.353	0.032	(0.291, 0.414)
0.8	0.362	0.032	(0.299, 0.424)
1	0.367	0.032	(0.305, 0.430)
1.2	0.371	0.032	(0.308, 0.434)
1.4	0.372	0.032	(0.309, 0.435)
1.6	0.372	0.032	(0.309, 0.435)

5. Discussion

For a square contingency table with the same ordered row and column classifications, we have proposed the measure to represent the degree of departure from the CS model based on the cumulative probability. The proposed measure $\Psi^{(\lambda)}$ is useful for comparing the degrees of departure from CS among different ordinal tables as shown in the example.

The measure $\Psi^{(\lambda)}$ can be reformulated using the power divergence [15] as follows.

$$\Psi^{(\lambda)} = \frac{\lambda(\lambda+1)}{2^{\lambda}-1} \cdot \frac{1}{2}\left[I^{(\lambda)}(G_{ij}^{U}; Q_{ij}) + I^{(\lambda)}(G_{ji}^{L}; Q_{ij}) \right] \qquad (\lambda > -1),$$

where

$$I^{(\lambda)}(a_{ij}; b_{ij}) = \frac{1}{\lambda(\lambda+1)} \sum\sum_{i<j} a_{ij} \left[\left(\frac{a_{ij}}{b_{ij}} \right)^{\lambda} - 1 \right].$$

When $\lambda = 0$, using the Kullback–Leibler (KL) divergence,

$$\Psi^{(0)} = \frac{1}{2\log 2}\left[I^{(0)}(G_{ij}^{U}; Q_{ij}) + I^{(0)}(G_{ji}^{L}; Q_{ij})\right],$$

where

$$I^{(0)}(a_{ij}; b_{ij}) = \sum\sum_{i<j} a_{ij}\left(\frac{a_{ij}}{b_{ij}}\right).$$

Especially when $\lambda = 0$, it can be easily seen that

$$\Psi^{(0)} = \frac{1}{2\log 2}\min_{W_{ij}}\left[I^{(0)}(G_{ij}^{U}; W_{ij}) + I^{(0)}(G_{ji}^{L}; W_{ij})\right],$$

where $\sum\sum_{i<j} W_{ij} = 1$ and $W_{ij} > 0$. Therefore, Q_{ij} in $\Psi^{(\lambda)}$ is the value of W_{ij} such that it minimizes the sum of the two KL divergences between $\{G_{ij}^{U}\}$ and $\{W_{ij}\}$, and $\{G_{ji}^{L}\}$ and $\{W_{ij}\}$ with the conditional asymmetric structure. Note that Q_{ij} does not minimize the power divergence for any λ ($\lambda > -1; \lambda \neq 0$).

The reader may be interested in how to select the value of the parameter λ in $\Psi^{(\lambda)}$ to compare different tables in a practical situation. It is recommended to compare the results using the various values of λ rather than to see the result from a specified value of λ. This is because it would be impossible to conclude which table has a larger departure from CS if the results differ depending on the value of λ (although the authors have not experienced such a case yet). On the other hand, it would be possible to draw a conclusion when the results agree for all the used values of λ. Thus, it seems to be safe to use various values of λ and to compare the obtained results. If the reader considers the interpretation of the measure to be important, adopting $\lambda = 0$ may be recommended from the discussion in the previous paragraph.

Consider the artificial cell and cumulative probability tables given in Table 3. This table has unspecified probabilities, p_{12} and p_{21}, where $p_{12} + p_{21} = 0.61$. In Table 3 (b), the ratios G_{13}/G_{31} and G_{23}/G_{32} are equal to 2. Figure 1 shows the value of measures $\Phi^{(0)}$ and $\Psi^{(0)}$ for different values of the ratio G_{12}/G_{21}. It can be seen from this figure that (1) the value of $\Psi^{(0)}$ takes 0 when $G_{12}/G_{21} = 2$, which is the same as G_{13}/G_{31} and G_{23}/G_{32}; namely, the CS model holds; (2) the value of $\Psi^{(0)}$ takes a larger value as G_{12}/G_{21} moves away from 2. The results for $\lambda \neq 0$ are similar and, thus, are not reported here.

From these results, the measure $\Psi^{(\lambda)}$ would be appropriate to represent the degree of departure from the CS model, because it is natural to consider that the departure from CS increases as G_{12}/G_{21} moves away from 2 in Table 3. Figure 1 shows that the measure $\Psi^{(0)}$ has a trend similar to the measure $\Phi^{(0)}$ in this artificial example.

Table 3. Artificial cell and cumulative probability tables; the cumulative probabilities in Table 3 (b) are calculated from the cell probabilities in Table 3 (a).

(**a**) Cell probability table.		
0.10	p_{12}	0.02
p_{21}	0.10	0.04
0.01	0.02	0.10
(**b**) Cumulative probability table.		
	$0.02 + p_{12}$	0.02
$0.01 + p_{21}$		0.06
0.01	0.03	

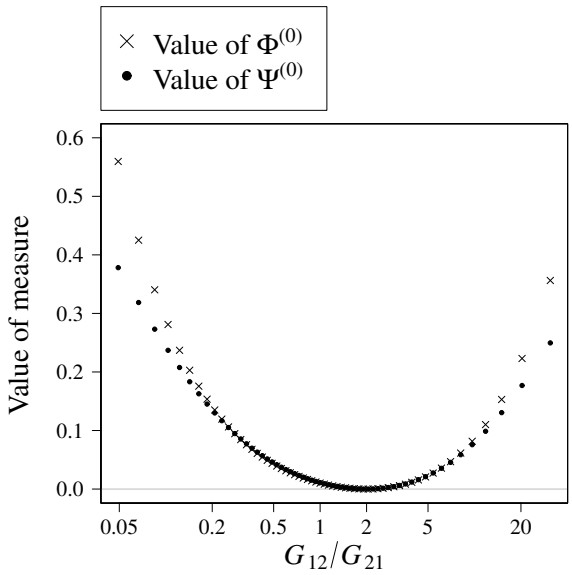

Figure 1. Values of measure $\Phi^{(0)}$ and $\Psi^{(0)}$ for values of G_{12}/G_{21} in Table 3.

We present another artificial example to clarify the difference between the two measures $\Phi^{(\lambda)}$ and $\Psi^{(\lambda)}$. Consider the artificial cell and cumulative probability tables given in Table 4. Figure 2 shows the values of the measures $\Phi^{(0)}$ and $\Psi^{(0)}$ for different values of the probability p_{13}. When $p_{13} = 0.1$, the ratios of the symmetric cell and cumulative probabilities are almost the same, and then the values of the two measures are nearly 0. When the value of p_{13} increases toward 0.5, the proportional structure of the cumulative probabilities is almost preserved.

Consequently, the value of $\Psi^{(0)}$ remains nearly 0. On the other hand, the value of $\Phi^{(0)}$ increases. This may be because the proportional structure of the cell probabilities is not preserved when the probability p_{13} increases. The results for $\lambda \neq 0$ are similar and, thus, are not reported here. Hence, the measure $\Psi^{(\lambda)}$ is appropriate to represent the degree of departure from a conditional symmetric structure of cumulative probabilities where the measure $\Phi^{(\lambda)}$ is not appropriate.

Table 4. Artificial cell and cumulative probability tables; the cumulative probabilities in Table 4 (b) are calculated from the cell probabilities in Table 4 (a).

(**a**) Cell probability table.		
0.05	0.02	p_{13}
0.10	0.05	0.02
$0.61 - p_{13}$	0.10	0.05
(**b**) Cumulative probability table.		
	$0.02 + p_{13}$	p_{13}
$0.71 - p_{13}$		$0.02 + p_{13}$
$0.61 - p_{13}$	$0.71 - p_{13}$	

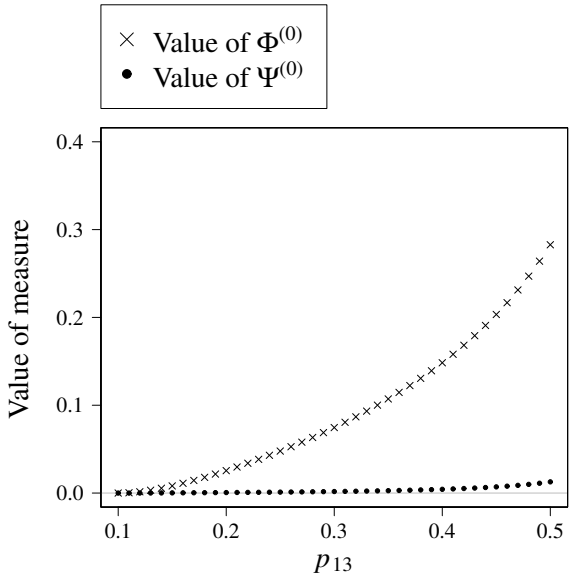

Figure 2. Values of measure $\Phi^{(0)}$ and $\Psi^{(0)}$ for values of p_{13} in Table 4.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the anonymous reviewers for their careful reading and their many insightful comments and suggestions.

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

Appendix A

To introduce the measure proposed by [11], we consider the reformulated restriction of the CS model using the conditional cell probabilities, that is,

$$p_{ij}^{U} = p_{ji}^{L} \quad (i < j),$$

where

$$p_{ij}^{U} = \Pr(X = i, Y = j | X < Y) = \frac{p_{ij}}{\delta_U},$$
$$p_{ji}^{L} = \Pr(X = j, Y = i | X > Y) = \frac{p_{ji}}{\delta_L},$$
$$\delta_U = \sum\sum_{s<t} p_{st}, \qquad \delta_L = \sum\sum_{s>t} p_{st}.$$

Assume that $\delta_U > 0, \delta_L > 0$, and $\{p_{ij} + p_{ji} > 0\}$. Let $q_{ij} = (p_{ij}^{U} + p_{ji}^{L})/2$ for $i < j$. The measure is defined by

$$\Phi^{(\lambda)} = \sum\sum_{i<j} q_{ij} \left[1 - \frac{\lambda 2^{\lambda}}{2^{\lambda} - 1} I_{ij}^{(\lambda)} \right] \quad (\lambda > -1),$$

where

$$I_{ij}^{(\lambda)} = \frac{1}{\lambda}\left[1 - (p_{ij}^c)^{\lambda+1} - (p_{ji}^c)^{\lambda+1}\right],$$
$$p_{ij}^c = \frac{p_{ij}^U}{p_{ij}^U + p_{ji}^L}, \qquad p_{ji}^c = \frac{p_{ji}^L}{p_{ij}^U + p_{ji}^L}.$$

Noting that $I_{ij}^{(\lambda)}$ is the diversity index of degree λ [14], this represents the degree to which the conditional probabilities (p_{ij}^c, p_{ji}^c) are distant from $(1/2, 1/2)$. The measure $\Psi^{(\lambda)}$ is the weighted mean of $I_{ij}^{(\lambda)}$ with the weights of q_{ij}. The value of the measure evaluated at $\lambda = 0$ is taken to be the limit as $\lambda \to 0$, namely,

$$\Phi^{(0)} = \sum\sum_{i<j} q_{ij}\left[1 - \frac{1}{\log 2} I_{ij}^{(0)}\right],$$

where

$$I_{ij}^{(0)} = -p_{ij}^c \log p_{ij}^c - p_{ji}^c \log p_{ji}^c.$$

Note that $I_{ij}^{(0)}$ is the Shannon entropy, which is a special case of the diversity index $I_{ij}^{(\lambda)}$.

References

1. Constantine, A.G.; Gower, J.C. Models for the analysis of interregional migration. *Environ. Plan. A* **1982**, *14*, 477–497. [CrossRef] [PubMed]
2. Agresti, A. *Categorical Data Analysis*, 3rd ed.; Wiley: New York, NY, USA, 2013. [CrossRef]
3. Bowker, A.H. A test for symmetry in contingency tables. *J. Am. Stat. Assoc.* **1948**, *43*, 572–574. [CrossRef] [PubMed]
4. Bishop, Y.M.M.; Fienberg, S.E.; Holland, P.W. *Discrete Multivariate Analysis: Theory and Practice*; The MIT Press: Cambridge, MA, USA, 1975. [CrossRef]
5. Tahata, K.; Tomizawa, S. Symmetry and asymmetry models and decompositions of models for contingency tables. *SUT J. Math.* **2014**, *50*, 131–165.
6. Gower, J.C. The analysis of asymmetry and orthogonality. In *Recent Developments in Statistics*; Barra, J.R., Ed.; Springer: Amsterdam, The Netherlands, 1977; pp. 109–123.
7. Constantine, A.G.; Gower, J.C. Graphical representation of asymmetric matrices. *Appl. Stat.* **1978**, *27*, 297–304. [CrossRef]
8. Tomizawa, S.; Murata, M. Gauss discrepancy type measure of degree of residuals from symmetry for square contingency tables. *J. Korean Stat. Soc.* **1992**, *21*, 59–69.
9. Greenacre, M. Correspondence analysis of square asymmetric matrices. *Appl. Stat.* **2000**, *49*, 297–310. [CrossRef]
10. McCullagh, P. A class of parametric models for the analysis of square contingency tables with ordered categories. *Biometrika* **1978**, *65*, 413–418. [CrossRef]
11. Tomizawa, S.; Saitoh, K. Measure of departure from conditional symmetry for square contingency tables with ordered categories. *J. Jpn. Stat. Soc.* **1999**, *29*, 65–78. [CrossRef]
12. Hashimoto, K. *Gendai Nihon No Kaikyu Kouzou (Class Structure in Modern Japan: Theory, Method and Quantitative Analysis)*; Toshindo Press: Tokyo, Japan, 1999. (In Japanese)
13. Iki, K.; Tahata, K.; Tomizawa, S. Measure of departure from marginal homogeneity using marginal odds for multi-way tables with ordered categories. *J. Appl. Stat.* **2012**, *39*, 279–295. [CrossRef]
14. Patil, G.P.; Taillie, C. Diversity as a concept and its measurement. *J. Am. Stat. Assoc.* **1982**, *77*, 548–561. [CrossRef]
15. Read, T.R.C.; Cressie, N. *Goodness-of-Fit Statistics for Discrete Multivariate Data*; Springer: New York, NY, USA, 1988. [CrossRef]