Austrian Journal of Statistics
January 2023, Volume 52, 62–71.
http://www.ajs.or.at/
doi:10.17713/ajs.v52i1.1382

An Index for the Degree and Directionality of Asymmetry for Square Contingency Tables with Ordered Categories

Tomotaka Momozaki
Tokyo University of Science

Tomoyuki Nakagawa
Tokyo University of Science

Kiyotaka Iki
Nihon University

Sadao Tomizawa
Tokyo University of Science

Abstract

For square contingency tables with ordered categories, an index based on Cressie and Read's power divergence (or Patil and Taillie's diversity index) has been proposed in order to measure the degree of departure from symmetry. Although there are two types of maximum asymmetry (i.e., whether (1) all the observations concentrate in the top-right cell in the table, or (2) they concentrate in the bottom-left cell), the existing index cannot distinguish the two directions of maximum asymmetry. This paper proposes a directional index based on an arc-cosine function in order to simultaneously represent the degree and directionality of asymmetry. The proposed index would be useful for comparing degrees of asymmetry for several square contingency tables. Numerical examples show the utility of the proposed index using some datasets. We evaluate the usefulness of the proposed index by applying it to real data of the clinical study. The proposed index provides analysis results that are easier to interpret than the existing index.

Keywords: arc-cosine function, asymmetry, contingency table, ordered category, symmetry.

1. Introduction

For two-way contingency tables, an analysis is generally performed to see whether the independence between the row and column classifications holds. On the other hand, for the analysis of square contingency tables with the same row and column classifications, there are many issues related to symmetry rather than independence. This is because, in square contingency tables, there is a strong association between the row and column classifications.

Consider an $r \times r$ square contingency table. Let p_{ij} denote the probability that an observation will fall in the ith row and jth column of the table ($i, j = 1, \ldots, r$). Bowker (1948) proposed the symmetry model defined by

$$p_{ij} = p_{ji} \text{ for all } i < j.$$

The symmetry model can be expressed as

$$U_{ij} = L_{ji} \text{ for all } i, j,$$

where

$$U_{ij} = \sum_{s=1}^{i} \sum_{t=j}^{r} p_{st}, \quad L_{ji} = \sum_{s=j}^{r} \sum_{t=1}^{i} p_{st}.$$

Note that the cumulative probabilities $\{U_{ij}, L_{ji}\}$ that may contain diagonal probabilities $\{p_{ii}\}$ can be defined for ordered categories.

The symmetry model, however, often does not hold when applied to real data. When the symmetry model fits real data poorly, other symmetry or asymmetry models (see, e.g., Tahata and Tomizawa 2014) are applied to real data.

We are also interested in measuring the degree of asymmetry when the symmetry model does not hold. When the symmetry model does not hold for multiple datasets, we may compare degrees of asymmetry. Over past years, many studies have proposed indexes to represent the degree of asymmetry. Tomizawa, Seo, and Yamamoto (1998) proposed an index based on power divergence (or diversity index) to represent the degree of departure from the symmetry model for square contingency tables with *nominal* categories. Moreover, Iki and Tomizawa (2018) proposed an index based on power divergence (or diversity index) to represent the degree of departure from the symmetry model for square contingency tables with *ordered* categories. For square contingency tables with *ordered* categories, we may be interested in distinguishing two types of maximum asymmetry (i.e., whether (1) all the observations concentrate in the top-right cell in the table, or (2) they concentrate in the bottom-left cell). The index of Iki and Tomizawa (2018), however, cannot distinguish the two directions of maximum asymmetry (see Appendix A.1 for the details of the index).

This paper proposes a directional index that can distinguish the two kinds of maximum asymmetry. Section 2 introduces the proposed index and shows the utility of the proposed index. Section 3 derives an approximate confidence interval for the proposed index. Section 4 shows the usefulness of the proposed index by applying it to real data of the clinical study. Section 5 describes the concluding remarks.

2. Directional index and its utility

This section proposes a directional index that can distinguish two types of maximum asymmetry and shows the utility of the proposed index.

2.1. Directional index via an arc-cosine function

Assuming that $p_{1r} + p_{r1} > 0$, we propose a directional index defined by

$$\Gamma = \frac{4}{\pi} \sum_{\substack{i,j \\ (i,j)\neq(r,1)}} (U^*_{ij} + L^*_{ji}) \left(\theta_{ij} - \frac{\pi}{4}\right),$$

where

$$U^*_{ij} = \frac{U_{ij}}{\tau}, \quad L^*_{ji} = \frac{L_{ji}}{\tau}, \quad \tau = \sum_{\substack{i,j \\ (i,j)\neq(r,1)}} (U_{ij} + L_{ji}), \quad \theta_{ij} = \arccos\left(\frac{U_{ij}}{\sqrt{U^2_{ij} + L^2_{ji}}}\right).$$

The range of Γ is -1 to 1 since the range of θ_{ij} is $0 \le \theta_{ij} \le \pi/2$. The index Γ has the following characteristics: (i) $\Gamma = -1$ if and only if $p_{1r} = 1$, (ii) $\Gamma = 1$ if and only if $p_{r1} = 1$; and (iii) if the symmetry model holds then $\Gamma = 0$.

Using the index Γ, we can see whether the degree of asymmetry departs toward the structure such that all the observations concentrate in the top-right cell $(1, r)$ in the table or the structure such that they concentrate in the bottom-left cell $(r, 1)$ in the table. As the index Γ

approaches -1, the asymmetry structure is closer to $p_{1r} = 1$, and as the index Γ approaches 1, it is closer to $p_{r1} = 1$.

Note that for all $\{i, j | (i, j) \neq (r, 1)\}$, $(\theta_{ij} - \pi/4)$ is zero when $U_{ij} = L_{ji}$, negative value when $U_{ij} > L_{ji}$, and positive value when $U_{ij} < L_{ji}$. Since the index Γ is the weighted sum of $(\theta_{ij} - \pi/4)$, the value of Γ is zero if and only if the weighted average of $(\theta_{ij} - \pi/4)$ is zero. If we shall refer to the structure of $\Gamma = 0$ as the average cumulative symmetry, the index Γ represents the degree of departure from the average cumulative symmetry towards the two types of maximum asymmetry.

2.2. The utility of the proposed index

This subsection demonstrates the utility of the proposed index. First, we compare the proposed index Γ with the indexes of Iki and Tomizawa (2018) (denoted by $\Phi^{(\lambda)}$; see Appendix A.1) and Tahata, Yamamoto, Nagatani, and Tomizawa (2009) (denoted by φ; see Appendix A.2). Tahata *et al.* (2009)'s index φ can distinguish whether (I) the complete upper asymmetry or (II) the complete lower asymmetry (i.e., whether (I) all the observations concentrate in the upper right triangle cells in the table, or (II) they concentrate in the lower left triangle cells). We consider the 4×4 structures of probability that have different asymmetric structures in Tables 1a to 1i. Table 2, however, represents that the values of $\Phi^{(1)}$ applied to Tables 1a and 1i, 1b and 1h, 1c and 1g, and 1d and 1f are equal. Additionally, it is impossible to calculate the values of φ applied to Tables 1a, 1b, 1h, and 1i because the existing index φ can be used under the condition $p_{ij} + p_{ji} > 0$ for all $i < j$. In contrast, all values of Γ are different, and the proposed index Γ can be applied to all of Tables 1a through 1i. When measuring the degree of asymmetry, it is important to distinguish between the two possible directions of maximum asymmetry, such that $p_{1r} = 1$ and $p_{r1} = 1$. This is because the interpretation of the result changes depending on whether the degree of asymmetry departs toward $p_{1r} = 1$ and $p_{r1} = 1$. It is also important to note that the value of the proposed index changes depending on the asymmetric structure. This is useful for comparing the degree of asymmetry among contingency tables. Thus, when measuring the degree of asymmetry, we suggest that analysts use the proposed index that can distinguish between the two directions of maximum asymmetry.

Table 1: The 4×4 structures of probability

(a)				(b)				(c)			
0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.3333	0.3334	0.1000	0.1000	0.1000	0.1000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.3333	0.0000	0.1000	0.1000	0.1000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1000	0.1000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1000
(d)				(e)				(f)			
0.0769	0.0769	0.0770	0.0770	0.0625	0.0625	0.0625	0.0625	0.0769	0.0769	0.0000	0.0000
0.0769	0.0769	0.0769	0.0770	0.0625	0.0625	0.0625	0.0625	0.0769	0.0769	0.0769	0.0000
0.0000	0.0769	0.0769	0.0769	0.0625	0.0625	0.0625	0.0625	0.0770	0.0769	0.0769	0.0769
0.0000	0.0000	0.0769	0.0769	0.0625	0.0625	0.0625	0.0625	0.0770	0.0770	0.0769	0.0769
(g)				(h)				(i)			
0.1000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.1000	0.1000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.1000	0.1000	0.1000	0.0000	0.3333	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.1000	0.1000	0.1000	0.1000	0.3334	0.3333	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000

Note that, in addition to the structure of probability where the symmetry model holds as shown in Table 1e, there is also a structure of probability where the value of the proposed index Γ is zero as shown in Table 3. Average cumulative symmetry includes such a structure of probability.

Table 2: The values of Γ, $\Phi^{(1)}$, and φ for Table 1

Index	Table number								
	1a	1b	1c	1d	1e	1f	1g	1h	1i
Γ	−1.000	−0.916	−0.561	−0.324	0.000	0.324	0.561	0.916	1.000
$\Phi^{(1)}$	1.000	0.846	0.324	0.113	0.000	0.113	0.324	0.846	1.000
φ	NA	NA	−1.000	−0.333	0.000	0.333	1.000	NA	NA

Table 3: The 4 × 4 structure of probability

0.150	0.025	0.050	0.025
0.050	0.150	0.025	0.025
0.025	0.025	0.150	0.050
0.025	0.050	0.025	0.150

3. Approximate confidence interval for the proposed index

Let n_{ij} denote the observed frequency in the ith row and jth column of the table ($i, j = 1, \dots, r$). Assume that a multinomial distribution applies to the $r \times r$ table. The sample proportions of $\{p_{ij}; i, j = 1, \dots, r\}$ are $\{\hat{p}_{ij} = n_{ij}/n\}$ with $n = \sum_{i,j} n_{ij}$. The estimator of the index Γ, $\hat{\Gamma}$ is provided by replacing $\{p_{ij}\}$ with $\{\hat{p}_{ij}\}$. This section derives the asymptotic distribution of $\hat{\Gamma}$ using the delta method (see, e.g., Agresti 2013), and the approximate confidence interval of Γ.

Let $\hat{\boldsymbol{p}}$ and $\boldsymbol{p}$ be the $r^2 \times 1$ vectors

$$\hat{\boldsymbol{p}} = (\hat{p}_{11}, \hat{p}_{12}, \dots, \hat{p}_{1r}, \hat{p}_{21}, \dots, \hat{p}_{rr})^\top \quad \text{and} \quad \boldsymbol{p} = (p_{11}, p_{12}, \dots, p_{1r}, p_{21}, \dots, p_{rr})^\top$$

respectively, where $\boldsymbol{a}^\top$ is the transpose of $\boldsymbol{a}$. Then $\sqrt{n}(\hat{\boldsymbol{p}} - \boldsymbol{p})$ has an asymptotically normal distribution with the mean zero vector and the covariance matrix $\mathrm{diag}(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^\top$, where $\mathrm{diag}(\boldsymbol{p})$ is a diagonal matrix with the elements of $\boldsymbol{p}$ on the main diagonal. Since

$$\hat{\Gamma} = \Gamma + d(\boldsymbol{p})(\hat{\boldsymbol{p}} - \boldsymbol{p}) + o_p(1)$$

with $d(\boldsymbol{p}) = \partial\Gamma/\partial\boldsymbol{p}^\top$, $\sqrt{n}(\hat{\Gamma} - \Gamma)$ asymptotically (as $n \to \infty$) has a normal distribution with the mean zero and the variance

$$\begin{aligned}\sigma^2[\Gamma] &= d(\boldsymbol{p})(\mathrm{diag}(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^\top)d(\boldsymbol{p})^\top \\ &= \sum_{k,l} p_{kl}\left(\frac{\partial\Gamma}{\partial p_{kl}}\right)^2 - \left(\sum_{k,l} p_{kl}\frac{\partial\Gamma}{\partial p_{kl}}\right)^2,\end{aligned}$$

where

$$\begin{aligned}\frac{\partial\Gamma}{\partial p_{kl}} =& \frac{4}{\pi\tau}\sum_{\substack{i,j \\ (i,j)\neq(r,1)}} \left(I[k \le i, j \le l] + I[j \le k, l \le i]\right)\left(\theta_{ij} - \frac{\pi}{4}(\Gamma + 1)\right) \\ &+ \frac{4}{\pi\tau}\sum_{\substack{i,j \\ (i,j)\neq(r,1)}} (U_{ij} + L_{ji})\frac{U_{ij}I[j \le k, l \le i] - L_{ji}I[k \le i, j \le l]}{U_{ij}^2 + L_{ji}^2}\end{aligned}$$

and $I[\cdot]$ is the function, $I[\cdot] = 1$ if true, 0 if not.

Let $\hat{\sigma}^2[\Gamma]$ denote $\sigma^2[\Gamma]$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$. The square root of $\hat{\sigma}^2[\Gamma]/n$ is an estimated standard error of $\hat{\Gamma}$, and

$$\hat{\Gamma} \pm z_{\alpha/2}\sqrt{\hat{\sigma}^2[\Gamma]/n}$$

is an approximate $100(1-\alpha)\%$ confidence interval for Γ, where $z_{\alpha/2}$ is the percentage point of the standard normal distribution corresponding to a two-tail probability of α.

4. Examples

Example 1: First, consider the data in Tables 4a and 4b, cited from Sugano, Kinoshita, Miwa, and Takeuchi (2012). In this study, 343 Japanese adult patients (aged $\geq$ 20 years) with a history of peptic ulcers were randomly assigned to treatment (esomeprazole, 175 patients; placebo, 168 patients). The modified LANZA score indicates that "0" is the best score and "+4" is the worst score. Thus, for the data in Tables 4a and 4b, the more all the observations concentrate in the top-right cell in the tables, the more improvement is shown. On the contrary, the more all the observations concentrate in the bottom-left cell in the tables, the more ingravescence is shown. As a matter of clinical interest, we would like to determine whether patients in the esomeprazole group improved more than patients in the placebo group. Therefore, for these data in Tables 4a and 4b, we are interested in comparing the degree of asymmetry as well as distinguishing the directionality for the two types of asymmetry. Table 5 represents that (i) the degree of asymmetry for the data in Table 4a departs toward the asymmetric structure, where all the observations concentrating in the top-right cell in the table since the confidence interval for Γ is negative, and (ii) the degree of asymmetry for the data in Table 4b departs toward the asymmetric structure, where all the observations concentrating in the bottom-left cell in the table since the confidence interval for Γ is positive. Thus, using the proposed index, we can see that patients in the esomeprazole group have experienced an improvement in terms of the change from baseline to endpoint in the modified LANZA score, and patients in the placebo group have experienced an ingravescence.

Note that the Tahata *et al.* (2009)'s index φ can be used only under the condition $p_{ij}+p_{ji}>0$ for all $i<j$. Therefore, it is impossible to calculate the values of φ applied to the data in Table 4 because $\hat{p}_{25}+\hat{p}_{52}=0$ for the data in the esomeprazole group and $\hat{p}_{14}+\hat{p}_{41}=0$ for the data in the placebo group. In contrast, the proposed index Γ can apply to the data if the data satisfy the only condition $\hat{p}_{1r}+\hat{p}_{r1}>0$. Comparing these conditions, $p_{1r}+p_{r1}>0$ and $p_{ij}+p_{ji}>0$ for all $i<j$, the proposed index Γ has high applicability and can be easily applied to the sparse data like the data in Table 4.

Table 4: Change in the modified LANZA score from baseline to end of study in (a) the esomeprazole group and (b) the placebo group; source Sugano *et al.* (2012)

End of study	Baseline					Total
	0	+1	+2	+3	+4	
(a) Esomeprazole group						
0	78	9	26	3	1	117
+1	1	5	6	4	0	16
+2	9	1	10	3	1	24
+3	1	0	1	0	0	2
+4	3	0	1	1	2	7
Total	92	15	44	11	4	166
(b) Placebo group						
0	41	2	19	0	0	62
+1	8	0	4	0	0	12
+2	12	4	14	3	0	33
+3	0	1	1	3	0	5
+4	29	7	11	6	0	53
Total	90	14	49	12	0	165

Table 5: Estimates of indexes Γ and $\Phi^{(1)}$, approximate standard errors for $\hat{\Gamma}$ and $\hat{\Phi}^{(1)}$, and approximate 95% confidence interval for Γ and $\Phi^{(1)}$ for the data in Tables 4a and 4b

Estimated index		Standard error	Confidence interval
(a) For Table 4a			
Γ	-0.198	0.056	$(-0.308, -0.088)$
$\Phi^{(1)}$	0.053	0.024	$(0.005, 0.101)$
(b) For Table 4b			
Γ	0.430	0.057	$(0.317, 0.542)$
$\Phi^{(1)}$	0.217	0.037	$(0.144, 0.291)$

Example 2: Next, consider the clinical trial data in Tables 6a and 6b, cited from Lundorff, Donnez, Korell, Audebert, Block, and DiZerega (2005). This study was a randomized (surgery plus Oxiplexw/Ap Gel group and surgery only group), the third party blinded, parallel-group design conducted at four centers in Europe. Patients were 18-46 years old requiring peritoneal cavity surgery by way of laparoscopy and expected to undergo a second-look laparoscopy as part of their treatment plan 6-10 weeks after the initial surgery. The American Fertility Society (AFS) adnexal adhesion score is determined by assessing the extent (area of adnexal organ covered by adhesions) and severity (severe: if the adhesion requires cutting to remove or tears peritoneal surfaces when removed bluntly or requires hemostasis; filmy if not severe) of the adhesions involving the Fallopian tube and ovary. Summing the scores for the Fallopian tube and the ovary provided a clinical category for the adhesion score: Minimum (0-5), Mild (6-10), Moderate (11-20), and Severe (21-32). Thus, for the data in Tables 6a and 6b, the more all the observations concentrate in the top-right cell in the tables, the more ingravescence is shown. On the contrary, the more all the observations concentrate in the bottom-left cell in the tables, the more improvement is shown. As a matter of clinical interest, we would like to determine whether patients in the surgery plus Oxiplexw/AP Gel group improved more than patients in the surgery-only group. Therefore, for these data in Tables 6a and 6b, we are interested in comparing the degree of asymmetry as well as distinguishing the directionality of the two types of asymmetry. Table 7 represents that (i) the degree of asymmetry for the data in Table 6a departs toward the asymmetric structure, where all the observations concentrating in the bottom-left cell in the table since the confidence interval for Γ is positive, and (ii) the degree of asymmetry for the data in Table 6b departs toward the asymmetric structure, where all the observations concentrating in the top-right cell in the table since the confidence interval for Γ is negative. Thus, using the proposed index, we can see that patients in the surgery plus Oxiplexw/AP Gel group have experienced an improvement in terms of the change from baseline to second-look in the AFS score, and patients in the surgery-only group have experienced an ingravescence.

Table 6: Change in the American Fertility Society (AFS) score category from baseline to second-look in (a) the surgery plus Oxiplexw/AP Gel group and (b) the surgery-only group; source Lundorff *et al.* (2005)

Baseline	Second-look				Total
	Minimal	Mild	Moderate	Severe	
(a) Surgery plus Oxiplexw/AP Gel group					
Minimal	22	1	0	0	23
Mild	2	2	1	0	5
Moderate	2	1	1	1	5
Severe	1	1	2	8	12
Total	27	5	4	9	45
(b) Surgery-only group					
Minimal	13	7	1	2	23
Mild	0	0	3	1	4
Moderate	0	0	1	4	5
Severe	0	0	1	8	9
Total	13	7	6	15	41

Table 7: Estimates of indexes Γ, $\Phi^{(1)}$, and φ, approximate standard errors for $\hat{\Gamma}$, $\hat{\Phi}^{(1)}$, and $\hat{\varphi}$, and approximate 95% confidence interval for Γ, $\Phi^{(1)}$, and φ for the data in Tables 6a and 6b

Estimated index		Standard error	Confidence interval
(a) For Table 6a			
Γ	0.195	0.081	$(0.036, 0.355)$
$\Phi^{(1)}$	0.063	0.043	$(-0.021, 0.147)$
φ	0.538	0.285	$(-0.022, 1.099)$
(b) For Table 6b			
Γ	-0.394	0.070	$(-0.532, -0.256)$
$\Phi^{(1)}$	0.206	0.054	$(0.100, 0.313)$
φ	-0.918	0.094	$(-1.101, -0.734)$

5. Concluding remarks

This paper proposed a directional index based on an arc-cosine function that can distinguish between the two directions of maximum asymmetry, namely, the structure such that all the observations concentrate in the top-right cell $(1, r)$ or the structure such that they concentrate in the bottom-left cell $(r, 1)$ in the table. Numerical examples demonstrated the utility of the proposed index by showing that the values of the proposed index for some asymmetric probability structures were all different, while the values of the existing index were partially the same. We recommend the proposed index for measuring the degree of asymmetry because the proposed index provides an easier interpretation for the analysis results than the existing index. Since the proposed index lies between -1 and 1 not depending on the number of categories r and the sample size n, the proposed index is useful to compare the degree of asymmetry for several square contingency tables. Note that the value of the proposed index Γ depends on the main diagonal cell probabilities $\{p_{ii}\}$, but does not the Tahata *et al.* (2009)'s index φ. Therefore, Γ rather than φ would be useful when one wants to utilize the information of the observations in the main diagonal cells of the table.

The proposed index should be applied to the tables with the ordered categories of the same row and column because the proposed index is not invariant under arbitrary similar permutations of the row and column categories.

Acknowledgements

The authors would like to thank the anonymous reviewer and the editor for careful reading and comments to improve this paper.

References

Agresti A (2013). *Categorical Data Analysis*. 3rd edition. John Wiley and Sons, Hoboken, New Jersey. ISBN 9781118710944.

Bowker AH (1948). "A Test for Symmetry in Contingency Tables." *Journal of the American Statistical Association*, **43**, 572–574. doi:10.1080/01621459.1948.10483284.

Iki K, Tomizawa S (2018). "Measure of Departure from Symmetry Using Cumulative Probability for Square Contingency Tables with Ordered Categories." *Advances and Applications in Statistics*, **52**, 401–412. doi:10.17654/AS052060401.

Lundorff P, Donnez J, Korell M, Audebert A, Block K, DiZerega G (2005). "Clinical Evaluation of a Viscoelastic Gel for Reduction of Adhesions Following Gynaecological Surgery by Laparoscopy in Europe." *Human Reproduction*, **20**, 514–520. doi:10.1093/humrep/deh651.

Sugano K, Kinoshita Y, Miwa H, Takeuchi T (2012). "Randomised Clinical Trial: Esomeprazole for the Prevention of Nonsteroidal Anti-Inflammatory Drug-Related Peptic Ulcers in Japanese Patients." *Alimentary Pharmacology and Therapeutics*, **36**, 115–125. doi:10.1111/j.1365-2036.2012.05133.x.

Tahata K, Tomizawa S (2014). "Symmetry and Asymmetry Models and Decompositions of Models for Contingency Tables." *SUT Journal of Mathematics*, **50**, 131–165. doi:10.20604/00000822.

Tahata K, Yamamoto K, Nagatani N, Tomizawa S (2009). "A Measure of Departure from Average Symmetry for Square Contingency Tables with Ordered Categories." *Austrian Journal of Statistics*, **38**, 101–108. doi:10.17713/ajs.v38i2.264.

Tomizawa S, Seo T, Yamamoto H (1998). "Power-Divergence-Type Measure of Departure from Symmetry for Square Contingency Tables that Have Nominal Categories." *Journal of Applied Statistics*, **25**, 387–398. doi:10.1080/02664769823115.

A. Appendix

A.1. Index based on the power divergence

Iki and Tomizawa (2018) proposed the index to represent the degree of departure from the symmetry model as follows. Assuming that $p_{1r} + p_{r1} > 0$, for $\lambda > -1$,

$$\Phi^{(\lambda)} = \frac{\lambda(\lambda+1)}{2^\lambda - 1} I^{(\lambda)},$$

where

$$I^{(\lambda)} = \frac{1}{\lambda(\lambda+1)} \sum_{\substack{i,j \\ (i,j)\neq(r,1)}} \left\{ U^*_{ij}\left[\left(\frac{U^*_{ij}}{W_{ij}}\right)^\lambda - 1\right] + L^*_{ji}\left[\left(\frac{L^*_{ji}}{W_{ij}}\right)^\lambda - 1\right]\right\},$$

with

$$U^*_{ij} = \frac{U_{ij}}{\tau}, \quad L^*_{ji} = \frac{L_{ji}}{\tau}, \quad U_{ij} = \sum_{s=1}^{i}\sum_{t=j}^{r} p_{st}, \quad L_{ji} = \sum_{s=j}^{r}\sum_{t=1}^{i} p_{st},$$

$$\tau = \sum_{\substack{i,j \\ (i,j)\neq(r,1)}} (U_{ij} + L_{ji}), \quad W_{ij} = \frac{U^*_{ij} + L^*_{ji}}{2}.$$

The $\Phi^{(0)}$ is defined as

$$\lim_{\lambda\to 0} \Phi^{(\lambda)} = \frac{1}{\log 2} I^{(0)},$$

where

$$I^{(0)} = \sum_{\substack{i,j \\ (i,j)\neq(r,1)}} \left(U^*_{ij} \log \frac{U^*_{ij}}{W_{ij}} + L^*_{ji} \log \frac{L^*_{ji}}{W_{ij}} \right).$$

The $I^{(\lambda)}$ is the power divergence between $\{U^*_{ij}, L^*_{ji}\}$ and $\{W_{ij}, W_{ji}\}$, and especially, $I^{(0)}$ is the Kullback Leibler information between them. The index $\Phi^{(\lambda)}$ has the following characteristics: (i) $0 \leq \Phi^{(\lambda)} \leq 1$; (ii) $\Phi^{(\lambda)} = 0$ if and only if the symmetry model holds; and (iii) $\Phi^{(\lambda)} = 1$ if and only if the degree of asymmetry is maximum in the sense that $p_{1r} = 1$ or $p_{r1} = 1$. However, index $\Phi^{(\lambda)}$ cannot distinguish between the two directions of maximum asymmetry. Namely, we cannot see whether the degree of asymmetry increases toward the structure such that all the observations concentrate in the top-right cell $(1, r)$ in the table or toward the structure such that all the observations concentrate in the bottom-left cell $(r, 1)$ in the table.

A.2. Index of complete asymmetry

Tahata *et al.* (2009) proposed the directional index to distinguish whether (I) the complete upper asymmetry or (II) the complete lower asymmetry (i.e., whether (I) all the observations concentrate in the upper right triangle cells in the table, or (II) they concentrate in the lower left triangle cells) as follows. Assuming that $p_{ij} + p_{ji} > 0$ for all $i < j$,

$$\varphi = \frac{4}{\pi} \sum_{i<j} (p^*_{ij} + p^*_{ji}) \left(\omega_{ij} - \frac{\pi}{4} \right)$$

where

$$p^*_{ij} = \frac{p_{ij}}{\delta}, \quad \delta = \sum_{i\neq j} p_{ij}, \quad \omega_{ij} = \arccos \left(\frac{p_{ij}}{\sqrt{p^2_{ij} + p^2_{ji}}} \right).$$

The directional index φ has the following characteristics: (i) $-1 \leq \varphi \leq 1$; (ii) $\varphi = -1$ if and only if $p_{ij} > 0$ for all $i < j$ (then $p_{ji} = 0$ for all $i < j$), say, complete-upper-asymmetry; (iii) $\varphi = 1$ if and only if $p_{ji} > 0$ for all $i < j$ (then $p_{ij} = 0$ for all $i < j$), say, complete-lower-asymmetry; and (iv) if the symmetry model holds then $\varphi = 0$.

Affiliation:

Tomotaka Momozaki
Department of Information Sciences
Graduate School of Science and Technology
Tokyo University of Science
Noda city, Chiba, Japan
E-mail: tomotaka.8823.momozaki@gmail.com

Tomoyuki Nakagawa
Department of Information Sciences
Faculty of Science and Technology
Tokyo University of Science
Noda city, Chiba, Japan
E-mail: t_nakagawa@rs.tus.ac.jp

Kiyotaka Iki
Department of Economics
Nihon University
Chiyoda-ku, Tokyo, Japan
E-mail: iki.kiyotaka@nihon-u.ac.jp

Sadao Tomizawa
Department of Information Sciences
Faculty of Science and Technology
Tokyo University of Science
Noda city, Chiba, Japan
E-mail: tomizawa@is.noda.tus.ac.jp

Austrian Journal of Statistics http://www.ajs.or.at/
published by the Austrian Society of Statistics http://www.osg.or.at/
Volume 52 *Submitted:* 2021-07-14
January 2023 *Accepted:* 2021-12-17