ATLANTIS PRESS

Journal of Statistical Theory and Applications
Vol. **19(2)**, June 2020, *pp.* 212–222
DOI: https://doi.org/10.2991/jsta.d.200507.001; ISSN 1538-7887
https://www.atlantis-press.com/journals/jsta

Measure of Departure from Marginal Homogeneity for the Analysis of Collapsed Square Contingency Tables with Ordered Categories

Kouji Yamamoto[1,*], Itsumi Iwama[2], Sadao Tomizawa[2]

[1] *Department of Biostatistics, Yokohama City University School of Medicine 3-9, Fukuura, Kanazawa-ku, Yokohama, Kanagawa, 236-0004, Japan*
[2] *Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science 2641, Yamazaki, Noda, Chiba, 278-8510, Japan*

ARTICLE INFO

Article History

Received 02 Feb 2019
Accepted 20 Oct 2019

Keywords

Collapsed table
Marginal homogeneity
Ordered category

2010 Mathematics Subject Classification: 62H17

ABSTRACT

For square contingency tables with ordered categories, there would be some situations that one would like to analyze them by using collapsed 3×3 tables combining some adjacent categories in the original table. This paper considers the marginal homogeneity for collapsed tables and proposes a measure which represents the degree of departure from the marginal homogeneity. The proposed measure lies between 0 and 1, and it takes zero when the marginal homogeneity holds. Examples are given.

1. INTRODUCTION

Consider an $r \times r$ square contingency table with the same row and column ordinal classifications. Let p_{ij} denote the probability that an observation will fall in the ith row and jth column of the table ($i = 1, \dots, r; j = 1, \dots, r$), and let X and Y denote the row and column variables, respectively. The marginal homogeneity (MH) model is defined by

$$Pr(X = i) = Pr(Y = i) \quad (i = 1, \dots, r),$$

namely

$$p_{i\cdot} = p_{\cdot i} \quad (i = 1, \dots, r),$$

where $p_{i\cdot} = \sum_{t=1}^{r} p_{it}$ and $p_{\cdot i} = \sum_{t=1}^{r} p_{ti}$ ([1], [2, p. 282]). This model is also expressed as

$$Pr(X = i | X \neq Y) = Pr(Y = i | X \neq Y) \quad (i = 1, \dots, r),$$

namely

$$p^c_{i\cdot} = p^c_{\cdot i} \quad (i = 1, \dots, r),$$

where

$$p^c_{i\cdot} = (p_{i\cdot} - p_{ii})/\delta, \quad p^c_{\cdot i} = (p_{\cdot i} - p_{ii})/\delta \quad \text{and } \delta = \sum_{i \neq j}\sum p_{ij}.$$

This indicates that the conditional row marginal distribution is identical to the conditional column marginal distribution under the condition that an observation will fall in one of the off-diagonal cells of the table.

*Corresponding author. Email: kouji_y@yokohama-cu.ac.jp

We consider the $(r-1)(r-2)/2$ (being $\binom{r-1}{2}$) ways of collapsing the $r \times r$ original table with ordered categories into a 3×3 table by choosing cut points after the sth and tth rows and after the sth and tth columns for $1 \le s < t \le r-1$. We define each collapsed 3×3 table as the T_{st} table $(1 \le s < t \le r-1)$. This process means that r categories in the $r \times r$ original table get into 3 categories (for example, A, B and C) by dividing at the cut points s and t and then a T_{st} table $(1 \le s < t \le r-1)$ has the new 3 categories. In a collapsed T_{st} table $(1 \le s < t \le r-1)$, let $G_{ij}^{(s,t)}$ indicate the corresponding probability for row value i $(i = 1, 2, 3)$ and column value j $(j = 1, 2, 3)$; that is,

$$G_{11}^{(s,t)} = \sum_{i=1}^{s}\sum_{j=1}^{s} p_{ij}, \quad G_{12}^{(s,t)} = \sum_{i=1}^{s}\sum_{j=s+1}^{t} p_{ij}, \quad G_{13}^{(s,t)} = \sum_{i=1}^{s}\sum_{j=t+1}^{r} p_{ij},$$
$$G_{21}^{(s,t)} = \sum_{i=s+1}^{t}\sum_{j=1}^{s} p_{ij}, \quad G_{22}^{(s,t)} = \sum_{i=s+1}^{t}\sum_{j=s+1}^{t} p_{ij}, \quad G_{23}^{(s,t)} = \sum_{i=s+1}^{t}\sum_{j=t+1}^{r} p_{ij},$$
$$G_{31}^{(s,t)} = \sum_{i=t+1}^{r}\sum_{j=1}^{s} p_{ij}, \quad G_{32}^{(s,t)} = \sum_{i=t+1}^{r}\sum_{j=s+1}^{t} p_{ij}, \quad G_{33}^{(s,t)} = \sum_{i=t+1}^{r}\sum_{j=t+1}^{r} p_{ij}.$$

Then, the MH model is expressed as

$$G_{i\cdot}^{(s,t)} = G_{\cdot i}^{(s,t)} \quad (i = 1, 2, 3),$$

where

$$G_{i\cdot}^{(s,t)} = \sum_{j=1}^{3} G_{ij}^{(s,t)}, \quad G_{\cdot i}^{(s,t)} = \sum_{j=1}^{3} G_{ji}^{(s,t)},$$

for all s and t $(1 \le s < t \le r-1)$; see [3].

When the MH model does not hold, we are interested in measuring the degree of departure from MH. For square contingency tables with nominal categories, Tomizawa and Makii [4] proposed a measure which expresses the degree of departure from MH. In addition, for square contingency tables with ordered categories, Tomizawa, Miyamoto and Ashihara [5] proposed the measure $\Gamma^{(\lambda)}$, which represents the degree of departure from MH. See Appendix A for the measure $\Gamma^{(\lambda)}$.

Moreover, when the MH does not hold, we are interested in measuring the degree of departure from MH for every collapsed table T_{st} $(1 \le s < t \le r-1)$. The purpose of this paper is to propose a new measure which expresses the degree of departure from MH by adopting collapsed 3 ×3 tables. Section 2 considers such a measure and Section 3 gives an approximate variance and a confidence interval for the measure. Section 4 shows examples using the proposed measure. Section 5 gives the concluding remarks.

2. MEASURE OF DEPARTURE FROM MH FOR COLLAPSED 3 × 3 TABLES

Let for $1 \le s < t \le r-1$,

$$\delta_{st} = \sum_{i=1}^{3} \sum_{\substack{j=1 \\ i \ne j}}^{3} G_{ij}^{(s,t)},$$

and

$$G_{i\cdot}^{c(s,t)} = \frac{1}{\delta_{st}}\left(G_{i\cdot}^{(s,t)} - G_{ii}^{(s,t)}\right), \quad G_{\cdot i}^{c(s,t)} = \frac{1}{\delta_{st}}\left(G_{\cdot i}^{(s,t)} - G_{ii}^{(s,t)}\right),$$

$$\pi_{st(i)}^{c*} = \frac{1}{2}\left(G_{i\cdot}^{c(s,t)} + G_{\cdot i}^{c(s,t)}\right) \quad (i = 1, 2, 3).$$

Assume that $\left\{G_{i\cdot}^{c(s,t)} + G_{\cdot i}^{c(s,t)}\right\}$ are all positive. The MH model is also represented as

$$G_{i\cdot}^{c(s,t)} = G_{\cdot i}^{c(s,t)} \left(= \pi_{st(i)}^{c*}\right) \quad (i = 1, 2, 3),$$

for $1 \le s < t \le r-1$.

Consider a measure defined by

$$\Omega^{(\lambda)} = \frac{1}{\binom{r-1}{2}} \sum\sum_{1\le s<t\le r-1} \Omega_{st}^{(\lambda)} \quad (\lambda > -1),$$

where

$$\Omega_{st}^{(\lambda)} = \frac{\lambda(\lambda+1)}{2(2^{\lambda}-1)} I_{st}^{(\lambda)} \left(\{G_{i\cdot}^{c(s,t)}, G_{\cdot i}^{c(s,t)}\}; \{\pi_{st(i)}^{c*}, \pi_{st(i)}^{c*}\}\right),$$

$$I_{st}^{(\lambda)}(\cdot;\cdot) = \frac{1}{\lambda(\lambda+1)} \sum_{i=1}^{3} \left[G_{i\cdot}^{c(s,t)} \left\{ \left(\frac{G_{i\cdot}^{c(s,t)}}{\pi_{st(i)}^{c*}} \right)^{\lambda} - 1 \right\} + G_{\cdot i}^{c(s,t)} \left\{ \left(\frac{G_{\cdot i}^{c(s,t)}}{\pi_{st(i)}^{c*}} \right)^{\lambda} - 1 \right\} \right],$$

and the value at $\lambda = 0$ is taken to be continuous limit as $\lambda \to 0$. Namely

$$\begin{aligned} \Omega^{(0)} &= \lim_{\lambda\to 0} \Omega^{(\lambda)} \\ &= \frac{1}{\binom{r-1}{2}} \sum\sum_{1\le s<t\le r-1} \Omega_{st}^{(0)}, \end{aligned}$$

where

$$\Omega_{st}^{(0)} = \frac{1}{2\log 2} I_{st}^{(0)} \left(\{G_{i\cdot}^{c(s,t)}, G_{\cdot i}^{c(s,t)}\}; \{\pi_{st(i)}^{c*}, \pi_{st(i)}^{c*}\}\right),$$

$$I_{st}^{(0)}(\cdot;\cdot) = \sum_{i=1}^{3} \left[G_{i\cdot}^{c(s,t)} \log\left(\frac{G_{i\cdot}^{c(s,t)}}{\pi_{st(i)}^{c*}} \right) + G_{\cdot i}^{c(s,t)} \log\left(\frac{G_{\cdot i}^{c(s,t)}}{\pi_{st(i)}^{c*}} \right) \right].$$

The submeasure $\Omega_{st}^{(\lambda)}$ $(1 \le s < t \le r-1)$ describes the degree of departure from the MH model for the collapsed T_{st} table. Note that $I_{st}^{(\lambda)}\left(\left\{G_{i\cdot}^{c(s,t)}, G_{\cdot i}^{c(s,t)}\right\}; \{\pi_{st(i)}^{c*}, \pi_{st(i)}^{c*}\}\right)$ is the power-divergence between $\left\{G_{i\cdot}^{c(s,t)}, G_{\cdot i}^{c(s,t)}\right\}$ and $\{\pi_{st(i)}^{c*}, \pi_{st(i)}^{c*}\}$, and especially, $I_{st}^{(0)}(\cdot)$ is the Kullback–Leibler information. Also, note that a real value λ (> -1) is chosen by users. See [6] for the power-divergence.

Let for $1 \le s < t \le r-1$,

$$G_{1(i)}^{c(s,t)} = \frac{G_{i\cdot}^{c(s,t)}}{G_{i\cdot}^{c(s,t)} + G_{\cdot i}^{c(s,t)}}, \quad G_{2(i)}^{c(s,t)} = \frac{G_{\cdot i}^{c(s,t)}}{G_{i\cdot}^{c(s,t)} + G_{\cdot i}^{c(s,t)}} \quad (i = 1, 2, 3).$$

Note that $\left\{G_{1(i)}^{c(s,t)} + G_{2(i)}^{c(s,t)} = 1\right\}$. The MH model may be expressed as

$$G_{1(i)}^{c(s,t)} = G_{2(i)}^{c(s,t)} \left(= \frac{1}{2}\right) \quad (i = 1, 2, 3),$$

for $1 \le s < t \le r-1$.

Then, $\Omega_{st}^{(\lambda)}$ $(1 \le s < t \le r-1)$ may also be defined by

$$\Omega_{st}^{(\lambda)} = \frac{\lambda(\lambda+1)}{2^{\lambda}-1} \sum_{i=1}^{3} \pi_{st(i)}^{c*} I_{st(i)}^{(\lambda)} \left(\left\{G_{k(i)}^{c(s,t)}\right\}; \left\{\frac{1}{2}\right\}\right) \quad (\lambda > -1),$$

where

$$I_{st(i)}^{(\lambda)}(\cdot;\cdot) = \frac{1}{\lambda(\lambda+1)} \left[G_{1(i)}^{c(s,t)} \left\{ \left(\frac{G_{1(i)}^{c(s,t)}}{1/2} \right)^{\lambda} - 1 \right\} + G_{2(i)}^{c(s,t)} \left\{ \left(\frac{G_{2(i)}^{c(s,t)}}{1/2} \right)^{\lambda} - 1 \right\} \right],$$

and the value at $\lambda = 0$ is taken to be continuous limit as $\lambda \to 0$. Namely,

$$\Omega_{st}^{(0)} = \frac{1}{\log 2} \sum_{i=1}^{3} \pi_{st(i)}^{c*} I_{st(i)}^{(0)} \left(\left\{G_{k(i)}^{c(s,t)}\right\}; \left\{\frac{1}{2}\right\}\right),$$

$$I_{st(i)}^{(0)}(\cdot;\cdot) = G_{1(i)}^{c(s,t)} \log\left(\frac{G_{1(i)}^{c(s,t)}}{1/2}\right) + G_{2(i)}^{c(s,t)} \log\left(\frac{G_{2(i)}^{c(s,t)}}{1/2}\right).$$

Therefore, $\Omega_{st}^{(\lambda)}$ represents the weighted sum of the power-divergence $I_{st(i)}^{(\lambda)}\left(\left\{G_{k(i)}^{c(s,t)}\right\};\left\{\frac{1}{2}\right\}\right)$. Moreover, $\Omega_{st}^{(\lambda)}$ can be expressed as

$$\Omega_{st}^{(\lambda)} = \sum_{i=1}^{3}\pi_{st(i)}^{c*}\left[1 - \frac{\lambda 2^{\lambda}}{2^{\lambda}-1} H_{st(i)}^{(\lambda)}\left(\left\{G_{k(i)}^{c(s,t)}\right\}\right)\right] \quad (\lambda > -1),$$

where

$$H_{st(i)}^{(\lambda)}\left(\left\{G_{k(i)}^{c(s,t)}\right\}\right) = \frac{1}{\lambda}\left(1 - \left(G_{1(i)}^{c(s,t)}\right)^{\lambda+1} - \left(G_{2(i)}^{c(s,t)}\right)^{\lambda+1}\right),$$

and the value at $\lambda = 0$ is taken to be continuous limit as $\lambda \to 0$. Namely

$$\Omega_{st}^{(0)} = \sum_{i=1}^{3}\pi_{st(i)}^{c*}\left[1 - \frac{1}{\log 2} H_{st(i)}^{(0)}\left(\left\{G_{k(i)}^{c(s,t)}\right\}\right)\right],$$

where

$$H_{st(i)}^{(0)}(\cdot) = -G_{1(i)}^{c(s,t)} \log G_{1(i)}^{c(s,t)} - G_{2(i)}^{c(s,t)} \log G_{2(i)}^{c(s,t)}.$$

Note that $H_{st(i)}^{(\lambda)}\left(\left\{G_{k(i)}^{(s,t)}\right\}\right)$ is Patil and Taillie's [7] diversity index of degree λ for $\left\{G_{1(i)}^{c(s,t)}, G_{2(i)}^{c(s,t)}\right\}$, which includes the Shannon entropy (when $\lambda = 0$). Therefore, $\Omega_{st}^{(\lambda)}$ represents the weighted sum of the diversity index $H_{st(i)}^{(\lambda)}\left(\left\{G_{k(i)}^{c(s,t)}\right\}\right)$. Note that for each $\lambda > -1$,

a. $0 \le H_{st(i)}^{(\lambda)}(\cdot) \le (2^{\lambda}-1)/(\lambda 2^{\lambda})$,

b. $H_{st(i)}^{(\lambda)}(\cdot) = 0$ if and only if $G_{1(i)}^{c(s,t)} = 1$ (then $G_{2(i)}^{c(s,t)} = 0$) or $G_{2(i)}^{c(s,t)} = 1$ (then $G_{1(i)}^{c(s,t)} = 0$),

c. $H_{st(i)}^{(\lambda)}(\cdot) = (2^{\lambda}-1)/(\lambda 2^{\lambda})$ if and only if $G_{1(i)}^{c(s,t)} = G_{2(i)}^{c(s,t)} \left(=\frac{1}{2}\right)$, that is $G_{i\cdot}^{c(s,t)} = G_{\cdot i}^{c(s,t)}$.

Thus, we conclude that the measure $\Omega^{(\lambda)}$ lies between 0 and 1, and the submeasure $\Omega_{st}^{(\lambda)}$ also lies between 0 and 1. For each $\lambda > -1$,

d. there is a structure of MH in the $r \times r$ table if and only if $\Omega^{(\lambda)} = 0$,

e. the degree of departure from MH in the $r \times r$ table is the largest, in the sense that $G_{1(i)}^{c(s,t)} = 1$ (then $G_{2(i)}^{c(s,t)} = 0$) or $G_{2(i)}^{c(s,t)} = 1$ (then $G_{1(i)}^{c(s,t)} = 0$) for $i = 1, 2, 3$ and $1 \le s < t \le r-1$ if and only if $\Omega^{(\lambda)} = 1$,

and for fixed s and t $(1 \le s < t \le r-1)$,

f. there is a structure of MH in a collapsed 3×3 table T_{st} if and only if $\Omega_{st}^{(\lambda)} = 0$,

g. the degree of departure from MH in a collapsed 3×3 table T_{st} is the largest, in the sense that $G_{1(i)}^{c(s,t)} = 1$ (then $G_{2(i)}^{c(s,t)} = 0$) or $G_{2(i)}^{c(s,t)} = 1$ (then $G_{1(i)}^{c(s,t)} = 0$) for $i = 1, 2, 3$ if and only if $\Omega_{st}^{(\lambda)} = 1$.

3. APPROXIMATE CONFIDENCE INTERVAL FOR MEASURE

Let n_{ij} denote the observed frequency in the ith row and jth column of the $r \times r$ table ($i = 1, \dots, r; j = 1, \dots, r$). Assuming that a multinomial distribution applies to the table, we shall consider an approximate standard error and a large sample confidence interval for the measure $\Omega^{(\lambda)}$, using the delta method. The sample version of $\Omega^{(\lambda)}$, that is, $\hat{\Omega}^{(\lambda)}$, is given by $\Omega^{(\lambda)}$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$, where $\hat{p}_{ij} = n_{ij}/n$ and $n = \sum\sum n_{ij}$. Using the delta method, $\sqrt{n}(\hat{\Omega}^{(\lambda)} - \Omega^{(\lambda)})$ has asymptotically (as $n \to \infty$) a normal distribution with mean 0 and variance $\sigma^2[\Omega^{(\lambda)}]$. See Appendix B for the details of $\sigma^2[\Omega^{(\lambda)}]$ and Appendix C for the details of the submeasure $\Omega_{st}^{(\lambda)}$.

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

4. EXAMPLES

Consider the data in Table 1. These data describe the cross-classification of father's and son's occupational status categories in Japan and British.

Since the confidence intervals for $\Omega^{(\lambda)}$ applied to the data in Tables 1a and 1b do not include zero for all λ (see Table 2), these would indicate that there is not the structure of MH in both tables.

When the degrees of departure from MH in Tables 1a and 1b are compared using the confidence intervals for $\Omega^{(\lambda)}$, it is greater in Table 1a than in Table 1b. Then, we can see that the degree of departure from MH is greater for Table 1a than for Table 1b. Therefore, the difference between the father's classification distribution and his son's distribution is greater in Japan than in British.

We shall further analyze the data in Tables 1a and 1b using the submeasure $\Omega_{st}^{(\lambda)}$ $(1 \le s < t \le r-1)$. We see from Table 3 that for Table 1a, (i) the degree of departure from MH in the collapsed table T_{12} is the smallest and (ii) those in the other collapsed tables are greater than that in the T_{12} table. Thus, it is seen that (i) when we combine the categories (3) to (5) in Table 1a, the degree of departure from MH for the collapsed table is slightest, and (ii) when we combine the categories in the other patterns in Table 1a, those for the collapsed tables are greater than that for case (i).

However, there is no possibility to decide in which collapsed table the degree of departure from MH is largest, because the values in the confidence intervals for them overlap each other except for $\Omega_{12}^{(\lambda)}$. So, we can get a conclusion, for Table 1a, that even if we get the original 5 categories into the 3 categories in any pattern to make us interpret the original table more easily, the marginal distributions for each

Table 1 | Cross-classification of father's and his son's social class in (a) Japan in 1975 [8, p. 151] and (b) British [2, p. 100].

		Son's Status					
	Father's Status	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**Total**
(a) Japan	(1)	29	43	25	31	4	132
	(2)	23	159	89	38	14	323
	(3)	11	69	184	34	10	308
	(4)	42	147	148	184	17	538
	(5)	42	176	377	114	298	1007
	Total	147	594	823	401	343	2308
(b) British	(1)	50	45	8	18	8	129
	(2)	28	174	84	154	55	495
	(3)	11	78	110	223	96	518
	(4)	14	150	185	714	447	1510
	(5)	3	42	72	320	411	848
	Total	106	489	459	1429	1017	3500

Table 2 | Estimates of measure $\Omega^{(\lambda)}$, approximate standard errors for $\hat{\Omega}^{(\lambda)}$ and approximate 95% confidence intervals for $\Omega^{(\lambda)}$, applied to Tables 1a and 1b.

Values of λ	$\hat{\Omega}^{(\lambda)}$	Standard Error	Confidence Interval
(a) For Table 1a			
−0.5	0.2332	0.0135	(0.2067, 0.2596)
0.0	0.3463	0.0176	(0.3119, 0.3807)
1.0	0.4226	0.0190	(0.3854, 0.4598)
1.5	0.4277	0.0190	(0.3904, 0.4649)
2.0	0.4226	0.0190	(0.3854, 0.4598)
(b) For Table 1b			
−0.5	0.0045	0.0016	(0.0013, 0.0076)
0.0	0.0075	0.0027	(0.0022, 0.0128)
1.0	0.0104	0.0037	(0.0031, 0.0176)
1.5	0.0106	0.0038	(0.0032, 0.0181)
2.0	0.0104	0.0037	(0.0031, 0.0176)

collapsed table are different between father and his son's social classes. Moreover, if we consider the original categories (1) as the high class, (2) as the middle class and (3) to (5) as the low class, social mobility between father and his son's social classes is less than that for the other classification patterns. We also see from Table 3 that for Table 1b, the confidence intervals for the submeasures $\Omega_{12}^{(\lambda)}$, $\Omega_{13}^{(\lambda)}$ and $\Omega_{23}^{(\lambda)}$ include zero, on the other hand, those for the other submeasures do not include zero. These results mean that when we combine the categories (3) to (5) in Table 1b, the marginal distribution for father's social class is equal to that for his son's one. Note that we can also obtain similar interpretations when we combine the categories (2) to (3) and (4) to (5), and (1) to (2) and (4) to (5) in Table 1b. Thus, if we consider the original categories (1) as the high class, (2) as the middle class and (3) to (5) as the low class, we could not see that social mobility between father's social class and his son's one happened. Also note that we are able to get a similar interpretations when we consider combining the categories (2) to (3) and (4) to (5), and (1) to (2) and (4) to (5) in Table 1b. As a result, for Table 1b, if we regard only the category (5) as the low class, when we consider the same specified category between father and his son, the probability that fathers whose son's social class is lower than him is different from the probability that sons whose father's social class is higher than him.

We can also see from Table 4 that the degree of departure from MH is greater for Table 1a than for Table 1b by using the measure $\Gamma^{(\lambda)}$ proposed by Tomizawa *et al.* [5]. However, when we would like to analyze these data in more detail, for example, by using collapsed tables as described above, it is impossible to do it by using the measure $\Gamma^{(\lambda)}$. So, in such a situation, the proposed measure $\Omega^{(\lambda)}$ would be useful.

Furthermore, we can know more about the degree of departure from MH in Tables 1a and 1b. From $\hat{\Omega}^{(\lambda)}$, for example, with $\lambda = 1$, (i) for Table 1a, the degree of departure from MH is estimated to be 42.26 percent of the maximum degree of departure from MH, and (ii) for Table 1b, the degree of departure from MH is estimated to be 1.04 percent of the maximum degree of departure from MH.

Table 3 | Estimates of submeasure $\Omega_{st}^{(\lambda)}$ applied to Tables 1a and 1b.

Estimated Submeasures	Values of λ	For Table 1a	Confidence Interval	For Table 1b	Confidence Interval
$\hat{\Omega}_{12}^{(\lambda)}$	−0.5	0.0731	(0.0491, 0.0971)	0.0016	(−0.0014, 0.0047)
	0.0	0.1200	(0.0818, 0.1581)	0.0028	(−0.0024, 0.0079)
	1.0	0.1608	(0.1113, 0.2103)	0.0038	(−0.0033, 0.0109)
	1.5	0.1643	(0.1139, 0.2146)	0.0039	(−0.0034, 0.0112)
	2.0	0.1608	(0.1113, 0.2103)	0.0038	(−0.0033, 0.0109)
$\hat{\Omega}_{13}^{(\lambda)}$	−0.5	0.2763	(0.2395, 0.3131)	0.0030	(−0.0008, 0.0067)
	0.0	0.4166	(0.3679, 0.4654)	0.0050	(−0.0013, 0.0114)
	1.0	0.5130	(0.4608, 0.5652)	0.0070	(−0.0018, 0.0157)
	1.5	0.5193	(0.4672, 0.5714)	0.0071	(−0.0019, 0.0161)
	2.0	0.5130	(0.4608, 0.5652)	0.0070	(−0.0018, 0.0157)
$\hat{\Omega}_{14}^{(\lambda)}$	−0.5	0.3327	(0.2893, 0.3762)	0.0092	(0.0025, 0.0159)
	0.0	0.4809	(0.4284, 0.5335)	0.0154	(0.0042, 0.0267)
	1.0	0.5720	(0.5194, 0.6246)	0.0213	(0.0058, 0.0368)
	1.5	0.5775	(0.5251, 0.6298)	0.0218	(0.0060, 0.0377)
	2.0	0.5720	(0.5194, 0.6246)	0.0213	(0.0058, 0.0368)
$\hat{\Omega}_{23}^{(\lambda)}$	−0.5	0.2211	(0.1915, 0.2507)	0.0024	(−0.0010, 0.0058)
	0.0	0.3394	(0.2985, 0.3803)	0.0041	(−0.0016, 0.0098)
	1.0	0.4256	(0.3796, 0.4717)	0.0056	(−0.0023, 0.0135)
	1.5	0.4317	(0.3855, 0.4779)	0.0058	(−0.0023, 0.0139)
	2.0	0.4256	(0.3796, 0.4717)	0.0056	(−0.0023, 0.0135)
$\hat{\Omega}_{24}^{(\lambda)}$	−0.5	0.2008	(0.1743, 0.2273)	0.0058	(0.0014, 0.0103)
	0.0	0.2925	(0.2596, 0.3255)	0.0098	(0.0023, 0.0173)
	1.0	0.3528	(0.3177, 0.3879)	0.0135	(0.0033, 0.0238)
	1.5	0.3569	(0.3217, 0.3920)	0.0139	(0.0033, 0.0244)
	2.0	0.3528	(0.3177, 0.3879)	0.0135	(0.0033, 0.0238)
$\hat{\Omega}_{34}^{(\lambda)}$	−0.5	0.2949	(0.2598, 0.3300)	0.0047	(0.0013, 0.0082)
	0.0	0.4282	(0.3848, 0.4716)	0.0080	(0.0022, 0.0137)
	1.0	0.5114	(0.4666, 0.5562)	0.0110	(0.0031, 0.0189)
	1.5	0.5165	(0.4718, 0.5612)	0.0113	(0.0031, 0.0194)
	2.0	0.5114	(0.4666, 0.5562)	0.0110	(0.0031, 0.0189)

5. CONCLUDING REMARKS

The measure $\Omega^{(\lambda)}$ always ranges from 0 to 1 independent of the dimension r and sample size n. Thus, it may be useful for comparing the degrees of departure from MH in several tables.

When the MH model does not hold, we are interested in (i) seeing what degree the departure from MH is for the original $r \times r$ table, (ii) seeing what degree the departure from MH is for T_{st} tables $(1 \le s < t \le r-1)$ and (iii) seeing in which T_{st} table $(1 \le s < t \le r-1)$ the degree of departure from MH is the largest. We recommend to use the proposed measure $\Omega^{(\lambda)}$ for (i) and the proposed submeasure $\Omega_{st}^{(\lambda)}$ for (ii) and (iii). The submeasure may also be used in the case that you would like to just analyze one collapsed table. Note that the measure $\Omega^{(\lambda)}$ is not invarianrt under the arbitrary permutations of row and column categories except the reverse order, so this measure should be applied only for ordinal data.

Consider the artificial data in Table 5. Let G^2 indicate the likelihood ratio statistic for goodness-of-fit of MH model. Table 6a gives the values of G^2 applied to these data. We shall compare the values of G^2 for Tables 5a and 5b. We see that the value of G^2 for Table 5a is greater than that for Table 5b. In contrast, for any fixed $\lambda(> -1)$, the value of $\hat{\Omega}^{(\lambda)}$ is greater for Table 5b than for Table 5a (see Table 6b). In terms of $\hat{G}_{i\cdot}^{c(s,t)} \big/ \hat{G}_{\cdot i}^{c(s,t)}$, $(i = 1, 2, 3,\ 1 \le s < t \le r-1)$ (see Table 5), it seems natural to conclude that the degree of departure from MH is less for Table 5a than for Table 5b. Therefore we recommend using $\hat{\Omega}^{(\lambda)}$ for comparing the degrees of departure from MH among several tables. To many readers, it might seem that G^2/n is also a reasonable measure for representing the degree of departure from MH. However, G^2/n is not such a measure for us. For instance, consider the artificial data in Tables 5b and 5c. The values of G^2/n are 0.0261 for Table 5b and 0.1284 for Table 5c. Therefore, the value of G^2/n is less for Table 5b than for Table 5c. On the other hand, for any fixed $\lambda(> -1)$, the value of $\hat{\Omega}^{(\lambda)}$ for Table 5b is equal to that for Table 5c (see Table 6b). Moreover, $\hat{G}_{i\cdot}^{c} \big/ \hat{G}_{\cdot i}^{c}$, $i = 1, 2, 3$ for Table 5b is identical to that for Table 5c (see Table 5). Therefore, it seems natural to get a conclusion that there are no differences between Tables 5b and 5c for the degree of departure from MH. As a result, $\hat{\Omega}^{(\lambda)}$ may also be more desirable to measure the degree of departure from MH than G^2/n.

Finally, the readers may be interested in considering a measure based on unconditional marginal probabilities (say, $\Gamma^{(\lambda)}$), instead of the proposed measure $\Omega^{(\lambda)}$ based on conditional marginal probabilities. The measure $\Gamma^{(\lambda)}$ takes the minimum value 0, when there is the structure of MH in the original table. On the other hand, we cannot define the maximum value, because it is impossible to define the structure of

Table 4 Estimates of measure $\Gamma^{(\lambda)}$ applied to Tables 1a and 1b.

Values of λ	For Table 1a	For Table 1b
−0.5	0.2730	0.0061
0.0	0.3990	0.0103
1.0	0.4799	0.0143
1.5	0.4850	0.0146
2.0	0.4799	0.0143

Table 5 Artificial data (n is sample size).

	(a) n = 2814				
	(1)	**(2)**	**(3)**	**(4)**	**Total**
(1)	251	266	37	42	596
(2)	140	329	271	98	838
(3)	72	76	224	189	561
(4)	32	20	310	457	819
Total	495	691	842	786	2814

Note:

$$\frac{\hat{G}_{1\cdot}^{c(1,2)}}{\hat{G}_{\cdot 1}^{c(1,2)}} = 1.41,\ \frac{\hat{G}_{2\cdot}^{c(1,2)}}{\hat{G}_{\cdot 2}^{c(1,2)}} = 1.41,\ \frac{\hat{G}_{3\cdot}^{c(1,2)}}{\hat{G}_{\cdot 3}^{c(1,2)}} = 0.45$$

$$\frac{\hat{G}_{1\cdot}^{c(1,3)}}{\hat{G}_{\cdot 1}^{c(1,3)}} = 1.41,\ \frac{\hat{G}_{2\cdot}^{c(1,3)}}{\hat{G}_{\cdot 2}^{c(1,3)}} = 0.79,\ \frac{\hat{G}_{3\cdot}^{c(1,3)}}{\hat{G}_{\cdot 3}^{c(1,3)}} = 1.10$$

$$\frac{\hat{G}_{1\cdot}^{c(2,3)}}{\hat{G}_{\cdot 1}^{c(2,3)}} = 2.24,\ \frac{\hat{G}_{2\cdot}^{c(2,3)}}{\hat{G}_{\cdot 2}^{c(2,3)}} = 0.55,\ \frac{\hat{G}_{3\cdot}^{c(2,3)}}{\hat{G}_{\cdot 3}^{c(2,3)}} = 1.10$$

(*continued*)

Table 5 Artificial data (n is sample size). *(Continued)*

(b) n = 2906

(1)	687	14	20	10	731
(2)	95	278	9	31	413
(3)	45	35	898	11	989
(4)	24	13	30	706	773
Total	851	340	957	758	2906

Note:

$$\frac{\hat{G}^{c(1,2)}_{1\cdot}}{\hat{G}^{c(1,2)}_{\cdot 1}} = 0.27, \frac{\hat{G}^{c(1,2)}_{2\cdot}}{\hat{G}^{c(1,2)}_{\cdot 2}} = 2.18, \frac{\hat{G}^{c(1,2)}_{3\cdot}}{\hat{G}^{c(1,2)}_{\cdot 3}} = 1.67$$

$$\frac{\hat{G}^{c(1,3)}_{1\cdot}}{\hat{G}^{c(1,3)}_{\cdot 1}} = 0.27, \frac{\hat{G}^{c(1,3)}_{2\cdot}}{\hat{G}^{c(1,3)}_{\cdot 2}} = 2.36, \frac{\hat{G}^{c(1,3)}_{3\cdot}}{\hat{G}^{c(1,3)}_{\cdot 3}} = 1.29$$

$$\frac{\hat{G}^{c(2,3)}_{1\cdot}}{\hat{G}^{c(2,3)}_{\cdot 1}} = 0.60, \frac{\hat{G}^{c(2,3)}_{2\cdot}}{\hat{G}^{c(2,3)}_{\cdot 2}} = 1.54, \frac{\hat{G}^{c(2,3)}_{3\cdot}}{\hat{G}^{c(2,3)}_{\cdot 3}} = 1.29$$

(c) n = 591

(1)	68	14	20	10	112
(2)	95	27	9	31	162
(3)	45	35	89	11	180
(4)	24	13	30	70	137
Total	232	89	148	122	591

Note:

$$\frac{\hat{G}^{c(1,2)}_{1\cdot}}{\hat{G}^{c(1,2)}_{\cdot 1}} = 0.27, \frac{\hat{G}^{c(1,2)}_{2\cdot}}{\hat{G}^{c(1,2)}_{\cdot 2}} = 2.18, \frac{\hat{G}^{c(1,2)}_{3\cdot}}{\hat{G}^{c(1,2)}_{\cdot 3}} = 1.67$$

$$\frac{\hat{G}^{c(1,3)}_{1\cdot}}{\hat{G}^{c(1,3)}_{\cdot 1}} = 0.27, \frac{\hat{G}^{c(1,3)}_{2\cdot}}{\hat{G}^{c(1,3)}_{\cdot 2}} = 2.36, \frac{\hat{G}^{c(1,3)}_{3\cdot}}{\hat{G}^{c(1,3)}_{\cdot 3}} = 1.29$$

$$\frac{\hat{G}^{c(2,3)}_{1\cdot}}{\hat{G}^{c(2,3)}_{\cdot 1}} = 0.60, \frac{\hat{G}^{c(2,3)}_{2\cdot}}{\hat{G}^{c(2,3)}_{\cdot 2}} = 1.54, \frac{\hat{G}^{c(2,3)}_{3\cdot}}{\hat{G}^{c(2,3)}_{\cdot 3}} = 1.29$$

Table 6 Values of G^2 and $\Omega^{(\lambda)}$ applied to Tables 5a, 5b and 5c.

		For Table 5a	For Table 5b	For Table 5c
(a) G^2				
		114.778	75.94	75.94
(b) $\hat{\Omega}^{(\lambda)}$				
	−0.5	0.0232	0.0653	0.0653
	0.0	0.0386	0.1060	0.1060
λ	1.0	0.0525	0.1405	0.1405
	1.5	0.0537	0.1434	0.1434
	2.0	0.0525	0.1405	0.1405

the furthest departure from MH. Thus, when the readers wants to consider the degree of departure from MH in addition to define the maximum departure from MH, we recommend using the measure $\Omega^{(\lambda)}$. If not, each measure, $\Gamma^{(\lambda)}$ and $\Omega^{(\lambda)}$, is enough to use.

CONFLICT OF INTEREST

Authors have no conflict of interest to declare.

REFERENCES

1. A. Stuart, Biometrika. 42 (1955), 412–416.
2. Y.M.M. Bishop, S.E. Fienberg, P.W. Holland, Discrete Multivariate Analysis: Theory and Practice, The MIT Press, Cambridge, MA, USA, 1975.
3. K. Yamamoto, K. Tahata, S. Tomizawa, Calcutta Stat. Assoc. Bull. 64 (2012), 21–36.
4. S. Tomizawa, T. Makii, J. Stat. Res. 35 (2001), 1–24.
5. S. Tomizawa, N. Miyamoto, N. Ashihara, Behaviormetrika. 30 (2003), 173–193.
6. T.R.C. Read, N. Cressie, Goodness-of-Fit Statistics for Discrete Multivariate Data, Springer, New York, NY, USA, 1988.
7. G.P. Patil, C. Taillie, J. Am. Stat. Assoc. 77 (1982), 548–561.
8. K. Hashimoto, Gendai Nihon no Kaikyuu Kouzou (Class Structure in Modern Japan: Theory, Method and Quantitative Analysis), Toshindo Press, Tokyo, Japan (in Japanese), 1999.

APPENDIX A

The measure to represent the degree of departure from MH proposed by Tomizawa *et al*. [5], is given as follows: assuming that $\{G_{1(i)} + G_{2(i)} \neq 0\}$,

$$\Gamma^{(\lambda)} = \frac{\lambda(\lambda+1)}{2^{\lambda}-1} I^{(\lambda)} \left(\left\{G^{*}_{1(i)}, G^{*}_{2(i)}\right\}; \left\{Q^{*}_{i}, Q^{*}_{i}\right\} \right) \quad (\lambda > -1),$$

where

$$I^{(\lambda)}(\cdot;\cdot) = \frac{1}{\lambda(\lambda+1)} \sum_{i=1}^{r-1} \left[G^{*}_{1(i)} \left\{ \left(\frac{G^{*}_{1(i)}}{Q^{*}_{i}} \right)^{\lambda} - 1 \right\} + G^{*}_{2(i)} \left\{ \left(\frac{G^{*}_{2(i)}}{Q^{*}_{i}} \right)^{\lambda} - 1 \right\} \right],$$

with

$$G_{1(i)} = \sum_{s=1}^{i} \sum_{t=i+1}^{r} p_{st}, \quad G_{2(i)} = \sum_{s=i+1}^{r} \sum_{t=1}^{i} p_{st}, \quad \Delta = \sum_{i=1}^{r-1} (G_{1(i)} + G_{2(i)}),$$

and

$$G^{*}_{1(i)} = \frac{G_{1(i)}}{\Delta}, \quad G^{*}_{2(i)} = \frac{G_{2(i)}}{\Delta}, \quad Q^{*}_{i} = \frac{1}{2} \left(G^{*}_{1(i)} + G^{*}_{2(i)} \right) \quad (i = 1, \ldots, r-1),$$

and the value at $\lambda = 0$ is taken to be continuous limit as $\lambda \to 0$.

APPENDIX B

Using the delta method, $\sqrt{n}(\hat{\Omega}^{(\lambda)} - \Omega^{(\lambda)})$ has aymptotically variance $\sigma^2[\Omega^{(\lambda)}]$ as follows:

$$\sigma^2 \left[\Omega^{(\lambda)} \right] = \sum_{k=1}^{r} \sum_{l=1}^{r} p_{kl} \left(\Delta^{(\lambda)}_{kl} \right)^2 \quad (\lambda > -1),$$

where

$$\Delta^{(\lambda)}_{kl} = \frac{1}{\binom{r-1}{2}} \sum_{s=1}^{r-2} \sum_{t=s+1}^{r-1} \frac{1}{\delta_{st}} \left(A^{(\lambda)(s,t)}_{kl} - \Omega^{(\lambda)}_{st} B^{(s,t)}_{kl} \right),$$

$$A^{(\lambda)(s,t)}_{kl} = \begin{cases} \sum_{i=1}^{3} \left[C_{kl(i)} \left\{ 1 - \frac{\lambda 2^{\lambda}}{2^{\lambda}-1} H^{(\lambda)}_{st(i)} \left(\left\{ G^{c(s,t)}_{k(i)} \right\} \right) \right\} + \delta_{st} \pi^{c*}_{st(i)} D^{(\lambda)}_{kl(i)} \right] & (\lambda \neq 0), \\ \sum_{i=1}^{3} \left[C_{kl(i)} \left\{ 1 - \frac{1}{\log 2} H^{(0)}_{st(i)} \left(\left\{ G^{c(s,t)}_{k(i)} \right\} \right) \right\} + \delta_{st} \pi^{c*}_{st(i)} D^{(0)}_{kl(i)} \right] & (\lambda = 0), \end{cases}$$

$$B_{kl} = 1 - \left(\sum_{i=1}^{3} E^{(s,t)}_{kl(i)} \right),$$

$$C^{(s,t)}_{kl(i)} = \frac{1}{2} \left(F^{(s,t)}_{kl(i)} + J^{(s,t)}_{kl(i)} - 2E^{(s,t)}_{kl(i)} \right),$$

$$D_{kl(i)}^{(\lambda)(s,t)} = \begin{cases} \dfrac{2^{\lambda}(\lambda+1)}{2(2^{\lambda}-1)} \dfrac{1}{\delta_{st}\pi_{st(i)}^{c*}} \left[\left(G_{1(i)}^{c(s,t)}\right)^{\lambda} \left\{ \left(F_{kl(i)}^{(s,t)} - E_{kl(i)}^{(s,t)}\right) - 2C_{kl(i)}^{(s,t)} G_{1(i)}^{c(s,t)} \right\} \right. \\ \qquad \left. + \left(G_{2(i)}^{c(s,t)}\right)^{\lambda} \left\{ \left(J_{kl(i)}^{(s,t)} - E_{kl(i)}^{(s,t)}\right) - 2C_{kl(i)}^{(s,t)} G_{2(i)}^{c(s,t)} \right\} \right] & (\lambda \neq 0), \\ \dfrac{1}{2\log 2} \dfrac{1}{\delta_{st}\pi_{st(i)}^{c*}} \left[\left\{ \log\left(G_{1(i)}^{c(s,t)}\right) + 1 \right\} \left\{ \left(F_{kl(i)}^{(s,t)} - E_{kl(i)}^{(s,t)}\right) - 2C_{kl(i)}^{(s,t)} G_{1(i)}^{c(s,t)} \right\} \right. \\ \qquad \left. + \left\{ \log\left(G_{2(i)}^{c(s,t)}\right) + 1 \right\} \left\{ \left(J_{kl(i)}^{(s,t)} - E_{kl(i)}^{(s,t)}\right) - 2C_{kl(i)}^{(s,t)} G_{2(i)}^{c(s,t)} \right\} \right] & (\lambda = 0), \end{cases}$$

$$E_{kl(i)}^{(s,t)} = \begin{cases} I(1 \leq k \leq s, 1 \leq l \leq s) & (i=1), \\ I(s+1 \leq k \leq t, s+1 \leq l \leq t) & (i=2), \\ I(t+1 \leq k \leq r, t+1 \leq l \leq r) & (i=3), \end{cases}$$

$$F_{kl(i)}^{(s,t)} = \begin{cases} I(1 \leq k \leq s) & (i=1), \\ I(s+1 \leq k \leq t) & (i=2), \\ I(t+1 \leq k \leq r) & (i=3), \end{cases}$$

$$J_{kl(i)}^{(s,t)} = \begin{cases} I(1 \leq l \leq s) & (i=1), \\ I(s+1 \leq l \leq t) & (i=2), \\ I(t+1 \leq l \leq r) & (i=3), \end{cases}$$

where $I(\cdot)$ is the indicator function.

APPENDIX C

Using the delta method, $\sqrt{n}\left(\hat{\Omega}_{st}^{(\lambda)} - \Omega_{st}^{(\lambda)}\right)$ $(1 \leq s < t \leq r-1)$ has aymptotically variance $\sigma^2\left[\Omega_{st}^{(\lambda)}\right]$ as follows:

$$\sigma^2\left[\Omega_{st}^{(\lambda)}\right] = \frac{1}{\delta_{st}^2} \left[\sum_{i=1}^{3} \sum_{\substack{j=1 \\ i \neq j}}^{3} G_{ij}^{(s,t)} \left(\eta_{ij}^{(\lambda)(s,t)}\right)^2 - \delta_{st}\left(\Omega_{st}^{(\lambda)}\right)^2 \right] \quad (\lambda > -1),$$

where

$$\eta_{ij}^{(\lambda)(s,t)} = \begin{cases} \dfrac{1}{2(2^{\lambda}-1)} \left[2^{\lambda} \left\{ \left(G_{1(i)}^{c(s,t)}\right)^{\lambda} + \lambda \left(G_{2(i)}^{c(s,t)}\right)^{\lambda} \left(\left(G_{1(i)}^{c(s,t)}\right)^{\lambda} - \left(G_{2(i)}^{c(s,t)}\right)^{\lambda} \right) \right. \right. \\ \qquad \left. \left. + \left(G_{2(j)}^{c(s,t)}\right)^{\lambda} - \lambda \left(G_{1(j)}^{c(s,t)}\right)^{\lambda} \left(\left(G_{1(j)}^{c(s,t)}\right)^{\lambda} - \left(G_{2(j)}^{c(s,t)}\right)^{\lambda} \right) \right\} - 2 \right] & (\lambda \neq 0), \\ 1 - \dfrac{1}{2\log 2} \left[-\log\left(G_{1(i)}^{c(s,t)}\right) - \log\left(G_{2(j)}^{c(s,t)}\right) \right] & (\lambda = 0). \end{cases}$$