Improved Estimator of Measure for Marginal Homogeneity using Marginal Odds in Square Contingency Tables

Kiyotaka Iki* and Sadao Tomizawa

Department of Information Sciences, Faculty of Science and Technology,
Tokyo University of Science, Noda City, Chiba, Japan
Email: iki@is.noda.tus.ac.jp

Abstract For square contingency tables, Iki, Tahata and Tomizawa (2011) considered the measure to represent the degree of departure from the marginal homogeneity model. Using the first-order term in the Taylor series expansion, the estimated measure with the cell probabilities replaced by the corresponding sample proportions is an approximately unbiased estimator when the sample size is large. The present paper proposes the improved approximate unbiased estimator of the measure which is obtained by using the second-order term in the Taylor series expansion. Also, it shows that the improved estimator approaches to the true measure faster than the original estimator as the sample size becomes larger by the simulation studies.

Keywords: Estimation, marginal homogeneity, marginal odds, square contingency table, taylor series expansion, unbiased estimator.

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 model is defined by

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

where $p_{i\cdot} = \sum_{t=1}^{R} p_{it}$ and $p_{\cdot i} = \sum_{s=1}^{R} p_{si}$; see Stuart (1955). This indicates that the row marginal distribution is identical to the column marginal distribution. This model is also expressed as

$$F_i^X = F_i^Y \quad (i = 1, \dots, R-1),$$

where $F_i^X = \sum_{k=1}^{i} p_{k\cdot}$ and $F_i^Y = \sum_{k=1}^{i} p_{\cdot k}$. Using the marginal logit, this model can be expressed as

$$L_i^X = L_i^Y \quad (i = 1, \dots, R-1),$$

where

$$L_i^X = \log\left(\frac{F_i^X}{1-F_i^X}\right), \quad L_i^Y = \log\left(\frac{F_i^Y}{1-F_i^Y}\right).$$

This states that the log odds that X is i or below instead of $i+1$ or above is equal to the log odds that Y is i or below instead of $i+1$ or above for $i = 1, \dots, R-1$. Further, the marginal homogeneity model is expressed as

$$H_{1(i)} = H_{2(i)} \quad (i = 1, \dots, R-1),$$

where

$$H_{1(i)} = \sum_{s=1}^{i} \sum_{t=i+1}^{R} p_{s\cdot} p_{\cdot t} = F_i^X (1 - F_i^Y),$$

$$H_{2(i)} = \sum_{s=i+1}^{R} \sum_{t=1}^{i} p_{s\cdot} p_{\cdot t} = (1 - F_i^X) F_i^Y.$$

This indicates that the probability that the row variable X selected at random from the row marginal distribution is in category i or below and the column variable Y selected independently at random from the column marginal distribution is in category $i+1$ or above is equal to the probability that such X is in category $i+1$ or above and such Y is in category i or below.

Since the marginal homogeneity model indicates that $\{H_{1(i)}\}$ are equal to corresponding $\{H_{2(i)}\}$, when the marginal homogeneity model does not hold, we are interested in a measure for seeing how far the probabilities $\{H_{1(i)}\}$ and $\{H_{2(i)}\}$ are distant from marginal homogeneity. Iki et al. (2011) considered the measure $\Phi^{(\lambda)}$ to represent the degree of departure from marginal homogeneity for the ordinal data, which is expressed by using the power-divergence (Read and Cressie, 1988, p. 15) or the Patil and Taillie's (1982) diversity index, and as a function of $\{H_{1(i)}\}$ and $\{H_{2(i)}\}$. Assuming that $\{H_{1(i)} + H_{2(i)} > 0\}$, let

$$\Delta = \sum_{m=1}^{R-1} \left(H_{1(m)} + H_{2(m)}\right),$$

and let

$$H^*_{1(i)} = \frac{H_{1(i)}}{\Delta}, \quad H^*_{2(i)} = \frac{H_{2(i)}}{\Delta}, \quad Q^*_i = \frac{1}{2}(H^*_{1(i)} + H^*_{2(i)}),$$
$$H^c_{1(i)} = \frac{H_{1(i)}}{H_{1(i)} + H_{2(i)}}, \quad H^c_{2(i)} = \frac{H_{2(i)}}{H_{1(i)} + H_{2(i)}} \quad (i = 1, \dots, R-1).$$

For $\lambda > -1$, the measure of departure from the marginal homogeneity model considered by Iki et al. (2011), is defined by

$$\begin{aligned}\Phi^{(\lambda)} &= \frac{1}{2^\lambda - 1} \sum_{i=1}^{R-1} \left[H^*_{1(i)} \left\{ \left(\frac{H^*_{1(i)}}{Q^*_i} \right)^\lambda - 1 \right\} + H^*_{2(i)} \left\{ \left(\frac{H^*_{2(i)}}{Q^*_i} \right)^\lambda - 1 \right\} \right] \\ &= 1 - \frac{2^\lambda}{2^\lambda - 1} \sum_{i=1}^{R-1} (H^*_{1(i)} + H^*_{2(i)}) \Big[1 - (H^c_{1(i)})^{\lambda+1} - (H^c_{2(i)})^{\lambda+1} \Big],\end{aligned}$$

and the value at $\lambda = 0$ is taken to be the limit as $\lambda \to 0$. The measure $\Phi^{(\lambda)}$ must lie between 0 and 1, and it would be useful for comparing the degrees of departure from marginal homogeneity toward the maximum departure in several tables.

Using the first-order term in the Taylor series expansion, the estimated measure with the cell probabilities replaced by the corresponding sample proportions is an approximately unbiased estimator when the sample size is large. Using the second-order term, Tahata et al. (2014) proposed the refined estimators of measures for marginal homogeneity proposed by Tomizawa and Makii (2001) and Tomizawa et al. (2003). So we are now interested in proposing the improved approximate unbiased estimator of $\Phi^{(\lambda)}$.

The purpose of the present paper is to propose the improved approximate unbiased estimator of $\Phi^{(\lambda)}$. Section 2 gives such a estimator. Section 3 shows that the proposed estimator works well in many cases by the simulation studies.

2 Improved Approximate Unbiased Estimator

Assume that the observed frequencies $\{n_{ij}\}$ have a multinomial distribution. Let p be the $R^2 \times 1$ probabilities vector

$$p = (p_{11}, p_{12}, \dots, p_{1R}, p_{21}, p_{22}, \dots, p_{2R}, \dots, p_{R1}, p_{R2}, \dots, p_{RR})^t,$$

whereĄgtĄhmeans transpose. Also let $\{\hat{p}_{ij}\}$ be the sample proportion, where $\hat{p}_{ij} = n_{ij}/n$ with $n = \sum\sum n_{ij}$ and let $\hat{p}$ be the $R^2 \times 1$ vector in the similar way. We assume that g has a nonzero differential at p, i.e., that g has the following expansion as $\hat{p} \to p$:

$$g(\hat{p}) = g(p) + \left[\frac{\partial g(p)}{\partial p^t} \right] (\hat{p} - p) + o(\|\hat{p} - p\|),$$

where $[\partial g(p)/\partial p^t]$ denotes $[\partial g(\hat{p})/\partial \hat{p}^t]$ evaluated at $\hat{p} = p$. For the details, see e.g., Agresti (2013, p. 589) and Bishop et al. (1975, p. 486). For large n, we can see from above equation that $g(\hat{p})$ is an approximate unbiased estimator of $g(p)$ because mean of $\hat{p}$ equals p. Similarly, the sample version of $\varPhi^{(\lambda)}$, i.e., $\hat{\varPhi}^{(\lambda)}$ is given by $\varPhi^{(\lambda)}$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$, is an asymptotically unbiased estimator of $\varPhi^{(\lambda)}$ when the sample size n is large.

Assuming that g has a second differential at p, $g(\hat{p})$ has the following expansion as $\hat{p} \to p$:

$$g(\hat{p}) = g(p) + \left[\frac{\partial g(p)}{\partial p^t}\right](\hat{p} - p) + \frac{1}{2}(\hat{p} - p)^t \left[\frac{\partial^2 g(p)}{\partial p \partial p^t}\right](\hat{p} - p) + o(\|\hat{p} - p\|^2),$$

where $[\partial^2 g(p)/\partial p \partial p^t]$ denotes $[\partial^2 g(\hat{p})/\partial \hat{p} \partial \hat{p}^t]$ evaluated at $\hat{p} = p$. Therefore when the sample size n is large, the mean of $g(\hat{p})$, i.e., $\mathrm{E}(g(\hat{p}))$, is approximately equal to

$$g(p) + \frac{1}{2n} tr\left(\left[\frac{\partial^2 g(p)}{\partial p \partial p^t}\right](D(p) - pp^t)\right),$$

where $D(p)$ denotes the $R^2 \times R^2$ diagonal matrix with the ith element of p as the ith diagonal element, because $\mathrm{Var}(\hat{p}) = \frac{1}{n}(D(p) - pp^t)$. Thus the mean of

$$g(\hat{p}) - \frac{1}{2n} tr\left(\left[\frac{\partial^2 g(p)}{\partial p \partial p^t}\right](D(p) - pp^t)\right)$$

is approximately equal to $g(p)$, and it would approach $g(p)$ faster than $g(\hat{p})$ as the sample size n becomes larger. However, since the second term is unknown, the improved estimator of $g(p)$ is given as follows:

$$g(\hat{p}) - \frac{1}{2n} tr\left(\left[\frac{\partial^2 g(\hat{p})}{\partial p \partial p^t}\right](D(\hat{p}) - \hat{p}\hat{p}^t)\right),$$

where $[\partial^2 g(\hat{p})/\partial p \partial p^t]$ is given by $[\partial^2 g(p)/\partial p \partial p^t]$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$ and $D(\hat{p})$ denotes $D(p)$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$.

We now propose the improved estimator of the true measure $\Phi^{(\lambda)}$ as follows:

$$\hat{\varPhi}^{(\lambda)*} = \hat{\varPhi}^{(\lambda)} - \frac{1}{2n} tr\left(\left[\frac{\partial^2 \hat{\varPhi}^{(\lambda)}}{\partial \hat{p} \partial \hat{p}^t}\right](D(\hat{p}) - \hat{p}\hat{p}^t)\right),$$

where $[\partial^2 \hat{\varPhi}^{(\lambda)}/\partial \hat{p} \partial \hat{p}^t]$ is given by $[\partial^2 \varPhi^{(\lambda)}/\partial p \partial p^t]$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$. Then, since $tr[\partial^2 \varPhi^{(\lambda)}/\partial p \partial p^t]pp^t = 0$, we note that

$$tr\left(\left[\frac{\partial^2 \varPhi^{(\lambda)}}{\partial p \partial p^t}\right](D(p) - pp^t)\right) = \sum_{k=1}^{R}\sum_{l=1}^{R} \frac{\partial^2 \varPhi^{(\lambda)}}{\partial p_{kl}^2} p_{kl},$$

where

$$\frac{\partial^2 \varPhi^{(\lambda)}}{\partial p_{kl}^2} = \frac{1 - \varPhi^{(\lambda)}}{\varDelta}\left(K_{2(kl)} - \frac{2}{\varDelta}\left(K_{1(kl)}\right)^2\right) + \frac{1}{\varDelta^2}\left\{2K_{1(kl)}L_{1(kl)}^{(\lambda)} - \varDelta L_{2(kl)}^{(\lambda)}\right\},$$

$$K_{1(kl)} = \sum_{m=1}^{R-1}\left(W_{1(kl)}^{(m)} + W_{2(kl)}^{(m)}\right),$$

$$K_{2(kl)} = \sum_{m=1}^{R-1}\left(W_{3(kl)}^{(m)} + W_{4(kl)}^{(m)}\right),$$

for $\lambda \neq 0$,

$$\begin{aligned}
L_{1(kl)}^{(\lambda)} = \frac{2^\lambda}{2^\lambda - 1} \sum_{i=1}^{R-1} \Bigg[& W_{1(kl)}^{(i)} \left\{ 1 - \left(H_{1(i)}^c\right)^\lambda - \lambda H_{2(i)}^c \left(\left(H_{1(i)}^c\right)^\lambda - \left(H_{2(i)}^c\right)^\lambda \right) \right\} \\
& + W_{2(kl)}^{(i)} \left\{ 1 - \left(H_{2(i)}^c\right)^\lambda - \lambda H_{1(i)}^c \left(\left(H_{2(i)}^c\right)^\lambda - \left(H_{1(i)}^c\right)^\lambda \right) \right\} \Bigg], \\
L_{2(kl)}^{(\lambda)} = \frac{2^\lambda}{2^\lambda - 1} \sum_{i=1}^{R-1} \Bigg[& W_{3(kl)}^{(i)} \left\{ 1 - \left(H_{1(i)}^c\right)^\lambda - \lambda H_{2(i)}^c \left(\left(H_{1(i)}^c\right)^\lambda - \left(H_{2(i)}^c\right)^\lambda \right) \right\} \\
& + W_{4(kl)}^{(i)} \left\{ 1 - \left(H_{2(i)}^c\right)^\lambda - \lambda H_{1(i)}^c \left(\left(H_{2(i)}^c\right)^\lambda - \left(H_{1(i)}^c\right)^\lambda \right) \right\} \\
& - \frac{\lambda(1+\lambda)}{H_{1(i)} + H_{2(i)}} \left(\left(H_{1(i)}^c\right)^{\lambda-1} + \left(H_{2(i)}^c\right)^{\lambda-1} \right) \left(W_{1(kl)}^{(i)} H_{2(i)}^c - W_{2(kl)}^{(i)} H_{1(i)}^c \right)^2 \Bigg],
\end{aligned}$$

and for $\lambda = 0$,

$$\begin{aligned}
L_{1(kl)}^{(0)} = \frac{1}{\log 2} \sum_{i=1}^{R-1} & \left(-W_{1(kl)}^{(i)} \log H_{1(i)}^c - W_{2(kl)}^{(i)} \log H_{2(i)}^c \right), \\
L_{2(kl)}^{(0)} = \frac{1}{\log 2} \sum_{i=1}^{R-1} & \Bigg\{ -W_{3(kl)}^{(i)} \log H_{1(i)}^c - W_{4(kl)}^{(i)} \log H_{2(i)}^c \\
& - \frac{1}{H_{1(i)}} W_{1(kl)}^{(i)} \left(W_{1(kl)}^{(i)} H_{2(i)}^c - W_{2(kl)}^{(i)} H_{1(i)}^c \right) \\
& - \frac{1}{H_{2(i)}} W_{2(kl)}^{(i)} \left(W_{2(kl)}^{(i)} H_{1(i)}^c - W_{1(kl)}^{(i)} H_{2(i)}^c \right) \Bigg\},
\end{aligned}$$

with

$$\begin{aligned}
W_{1(kl)}^{(m)} &= \sum_{a=1}^{m} \sum_{b=m+1}^{R} \left(I_{(a=k)} p_{\cdot b} + p_{a\cdot} I_{(b=l)} \right), \\
W_{2(kl)}^{(m)} &= \sum_{a=m+1}^{R} \sum_{b=1}^{m} \left(I_{(a=k)} p_{\cdot b} + p_{a\cdot} I_{(b=l)} \right), \\
W_{3(kl)}^{(m)} &= \sum_{a=1}^{m} \sum_{b=m+1}^{R} \left(I_{(a=k)} I_{(b=l)} + I_{(a=k)} I_{(b=l)} \right), \\
W_{4(kl)}^{(m)} &= \sum_{a=m+1}^{R} \sum_{b=1}^{m} \left(I_{(a=k)} I_{(b=l)} + I_{(a=k)} I_{(b=l)} \right),
\end{aligned}$$

and where $I_{(\cdot)}$ is the indicator function, $I_{(\cdot)} = 1$ if true, 0 if not. Therefore, the improved estimator $\hat{\Phi}^{(\lambda)*}$ is also expressed as follows:

$$\hat{\Phi}^{(\lambda)*} = \hat{\Phi}^{(\lambda)} - \frac{1}{2n} \sum_{k=1}^{R} \sum_{l=1}^{R} \frac{\partial^2 \hat{\Phi}^{(\lambda)}}{\partial \hat{p}_{kl}^2} \hat{p}_{kl}.$$

3 Simulation Studies

By the simulation studies, we calculate the values of estimated measures $\hat{\Phi}^{(\lambda)}$ and $\hat{\Phi}^{(\lambda)*}$ from the observed frequencies of sample size $n = 30$, 40, 50, 100, 500 and 1000, which are obtained from the true probability distribution (see Tables 1a to 6a). We shall compare the mean of the values of $\hat{\Phi}^{(\lambda)}$ and $\hat{\Phi}^{(\lambda)*}$ obtained by 1000 times simulations, for each sample size. The results of simulations are given in Tables 1c to 6c.

Tables 1a, 3a and 5a have a characteristic that the sum of the probabilities of main-diagonal cells is very small ($p_{ii} = 0.020$ for $i = 1, 2, 3, 4$) and Tables 2a, 4a and 6a have a characteristic that the sum of the probabilities of main-diagonal cells is large ($p_{ii} = 0.100$ for $i = 1, 2, 3, 4$). Also the true values of measures for Tables 1a and 2a are small, while those for Tables 3a and 4a are medium, and those for Tables 5a and 6a are large, respectively.

We can see that the improved estimator $\hat{\Phi}^{(\lambda)*}$ approaches the true value $\Phi^{(\lambda)}$ faster than the original estimator $\hat{\Phi}^{(\lambda)}$ when $\lambda \geq 1$ from Tables 1c to 6c. Especially, we can see great improvement when sample size is small.

Table 1. (a) The artificial probabilities $\{p_{ij}\}$, (b) the value of $\Phi^{(\lambda)}$ and (c) the mean of the values of estimated measures obtained by generating 1000 times simulations, with each sample size n, for Table 1a.

(a)

	(1)	(2)	(3)	(4)
(1)	0.020	0.101	0.085	0.038
(2)	0.066	0.020	0.123	0.140
(3)	0.042	0.063	0.020	0.110
(4)	0.040	0.051	0.061	0.020

(b)

λ	$\Phi^{(\lambda)}$
1.0	0.1118
3.0	0.0978

(c)

λ	n	$\hat{\Phi}^{(\lambda)}$	$\hat{\Phi}^{(\lambda)*}$
1.0	30	0.1940	0.1286
	40	0.1660	0.1144
	50	0.1581	0.1162
	100	0.1330	0.1112
	500	0.1146	0.1101
	1000	0.1138	0.1115
3.0	30	0.1677	0.1028
	40	0.1533	0.1041
	50	0.1415	0.1013
	100	0.1185	0.0978
	500	0.1001	0.0959
	1000	0.1005	0.0984

Table 2. (a) The artificial probabilities $\{p_{ij}\}$, (b) the value of $\Phi^{(\lambda)}$ and (c) the mean of the values of estimated measures obtained by generating 1000 times simulations, with each sample size n, for Table 2a.

(a)

	(1)	(2)	(3)	(4)
(1)	0.100	0.052	0.068	0.110
(2)	0.044	0.100	0.054	0.058
(3)	0.038	0.042	0.100	0.052
(4)	0.020	0.020	0.042	0.100

(b)

λ	$\Phi^{(\lambda)}$
1.0	0.1129
3.0	0.0986

(c)

λ	n	$\hat{\Phi}^{(\lambda)}$	$\hat{\Phi}^{(\lambda)*}$
1.0	30	0.1614	0.1153
	40	0.1492	0.1140
	50	0.1431	0.1149
	100	0.1318	0.1175
	500	0.1154	0.1125
	1000	0.1141	0.1127
3.0	30	0.1504	0.1062
	40	0.1302	0.0960
	50	0.1234	0.0961
	100	0.1118	0.0979
	500	0.1016	0.0988
	1000	0.1014	0.1000

4 Concluding Remarks

The present paper has proposed the improved approximate unbiased estimator $\hat{\Phi}^{(\lambda)*}$ of the true measure $\Phi^{(\lambda)}$ proposed by Iki et al. (2011), however, Tahata et al. (2014) proposed the refined estimators of measures proposed by Tomizawa and Makii (2001) and Tomizawa et al. (2003).

Table 3. (a) The artificial probabilities $\{p_{ij}\}$, (b) the value of $\Phi^{(\lambda)}$ and (c) the mean of the values of estimated measures obtained by generating 1000 times simulations, with each sample size n, for Table 3a.

(a)

	(1)	(2)	(3)	(4)
(1)	0.020	0.096	0.100	0.068
(2)	0.041	0.020	0.163	0.170
(3)	0.027	0.023	0.020	0.160
(4)	0.040	0.011	0.021	0.020

(b)

λ	$\Phi^{(\lambda)}$
1.0	0.4391
3.0	0.4066

(c)

λ	n	$\hat{\Phi}^{(\lambda)}$	$\hat{\Phi}^{(\lambda)*}$
1.0	30	0.4669	0.4371
	40	0.4501	0.4272
	50	0.4556	0.4381
	100	0.4515	0.4431
	500	0.4410	0.4393
	1000	0.4398	0.4390
3.0	30	0.4375	0.4035
	40	0.4319	0.4065
	50	0.4250	0.4044
	100	0.4193	0.4092
	500	0.4091	0.4070
	1000	0.4049	0.4039

Table 4. (a) The artificial probabilities $\{p_{ij}\}$, (b) the value of $\Phi^{(\lambda)}$ and (c) the mean of the values of estimated measures obtained by generating 1000 times simulations, with each sample size n, for Table 4a.

(a)

	(1)	(2)	(3)	(4)
(1)	0.100	0.052	0.128	0.135
(2)	0.014	0.100	0.079	0.088
(3)	0.008	0.012	0.100	0.062
(4)	0.005	0.005	0.012	0.100

(b)

λ	$\Phi^{(\lambda)}$
1.0	0.4430
3.0	0.4079

(c)

λ	n	$\hat{\Phi}^{(\lambda)}$	$\hat{\Phi}^{(\lambda)*}$
1.0	30	0.4670	0.4465
	40	0.4571	0.4416
	50	0.4562	0.4439
	100	0.4490	0.4428
	500	0.4442	0.4429
	1000	0.4438	0.4432
3.0	30	0.4386	0.4141
	40	0.4186	0.3997
	50	0.4194	0.4041
	100	0.4122	0.4045
	500	0.4102	0.4086
	1000	0.4089	0.4081

Table 5. (a) The artificial probabilities $\{p_{ij}\}$, (b) the value of $\Phi^{(\lambda)}$ and (c) the mean of the values of estimated measures obtained by generating 1000 times simulations, with each sample size n, for Table 5a.

(a)

	(1)	(2)	(3)	(4)
(1)	0.020	0.181	0.105	0.202
(2)	0.006	0.020	0.133	0.110
(3)	0.002	0.013	0.020	0.143
(4)	0.003	0.011	0.011	0.020

(b)

λ	$\Phi^{(\lambda)}$
1.0	0.7797
3.0	0.7564

(c)

λ	n	$\hat{\Phi}^{(\lambda)}$	$\hat{\Phi}^{(\lambda)*}$
1.0	30	0.7830	0.7787
	40	0.7814	0.7786
	50	0.7814	0.7792
	100	0.7817	0.7808
	500	0.7799	0.7798
	1000	0.7803	0.7802
3.0	30	0.7693	0.7623
	40	0.7590	0.7538
	50	0.7580	0.7539
	100	0.7573	0.7554
	500	0.7562	0.7559
	1000	0.7563	0.7561

Table 6. (a) The artificial probabilities $\{p_{ij}\}$, (b) the value of $\Phi^{(\lambda)}$ and (c) the mean of the values of estimated measures obtained by generating 1000 times simulations, with each sample size n, for Table 6a.

(a)

	(1)	(2)	(3)	(4)
(1)	0.100	0.002	0.008	0.559
(2)	0.004	0.100	0.004	0.006
(3)	0.008	0.002	0.100	0.002
(4)	0.001	0.002	0.002	0.100

(b)

λ	$\Phi^{(\lambda)}$
1.0	0.7662
3.0	0.7407

(c)

λ	n	$\hat{\Phi}^{(\lambda)}$	$\hat{\Phi}^{(\lambda)*}$
1.0	30	0.7647	0.7638
	40	0.7664	0.7660
	50	0.7675	0.7673
	100	0.7664	0.7664
	500	0.7663	0.7766
	1000	0.7663	0.7663
3.0	30	0.7473	0.7445
	40	0.7405	0.7383
	50	0.7417	0.7401
	100	0.7402	0.7395
	500	0.7411	0.7410
	1000	0.7399	0.7399

From the simulation studies, we conclude that the improved estimator $\hat{\Phi}^{(\lambda)*}$ tends to approach to the true value $\Phi^{(\lambda)}$ faster than the estimator $\hat{\Phi}^{(\lambda)}$ as the sample size n becomes larger, when $\lambda \geq 1$.

When $\lambda < 1$, we can calculate the improved estimator $\hat{\Phi}^{(\lambda)*}$ for only the case of $H_{1(i)} > 0$ and $H_{2(i)} > 0$ for $i = 1, \ldots, R-1$, i.e., $p_{1\cdot} > 0$, $p_{R\cdot} > 0$, $p_{\cdot 1} > 0$ and $p_{\cdot R} > 0$. On the other hand, the original estimator $\hat{\Phi}^{(\lambda)}$ can be calculated for the case of $H_{1(i)} + H_{2(i)} > 0$ for $i = 1, \ldots, R-1$. In other words, the calculable conditions are different between the improved estimator and the original estimator. Thus, it seems difficult to evaluate whether the improved estimator tends to approach the true value faster than the original estimator by simulation study when $\lambda < 1$. Therefore, we recommend that the proposed estimator should be used for the case of $\lambda \geq 1$. Then this estimator works very well.

Acknowledgments. The authors would like to thank the referee for their helpful comments.

References

1. Agresti, A. (2013). *Categorical Data Analysis,* 3rd ed. New Jersey: Wiley.
2. Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice.* Cambridge: The MIT Press.
3. Iki, K., Tahata, K. and Tomizawa, S. (2011). Measure of departure from marginal homogeneity using marginal odds for multi-way tables with ordered categories. *Journal of Applied Statistics*, Vol. 39, pp. 279-295.
4. Patil, G. P. and Taillie, C. (1982). Diversity as a concept and its measurement. *Journal of the American Statistical Association*, Vol. 77, pp. 548-561.
5. Read, T. R. C. and Cressie, N. (1988). *Goodness-of-Fit Statistics for Discrete Multivariate Data.* New York: Springer-Verlag.
6. Stuart, A. (1955). A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika*, Vol. 42, pp. 412-416.
7. Tahata, K., Tanaka, H. and Tomizawa, S. (2014). Refined estimators of measures for marginal homogeneity in square contingency tables. *International Journal of Pure and Applied Mathematics*, Vol. 90, pp. 501-513.
8. Tomizawa, S. and Makii, T. (2001). Generalized measures of departure from marginal homogeneity for contingency tables with nominal categories. *Journal of Statistical Research*, Vol. 35, pp. 1-24.
9. Tomizawa, S., Miyamoto, N. and Ashihara, N. (2003). Measure of departure from marginal homogeneity for square contingency tables having ordered categories. *Behaviormetrika*, Vol. 30, pp. 173-193.

