International Journal of Pure and Applied Mathematics
Volume 90 No. 4 2014, 501-513
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
doi: http://dx.doi.org/10.12732/ijpam.v90i4.11

ijpam.eu

REFINED ESTIMATORS OF MEASURES FOR MARGINAL HOMOGENEITY IN SQUARE CONTINGENCY TABLES

Kouji Tahata[1 §], Hiroyuki Tanaka[2], Sadao Tomizawa[3]
[1,2,3]Department of Information Sciences
Faculty of Science and Technology
Tokyo University of Science
2641 Yamazaki, Noda City, Chiba, 278-8510, JAPAN

Abstract: For square contingency tables, Tomizawa [6], Tomizawa and Makii [7] and Tomizawa, Miyamoto and Ashihara [8] considered the measures to represent the degree of departure from the marginal homogeneity model.

Using the first-order term in the Taylor series expansion, the estimated measures with the cell probabilities replaced by the corresponding sample proportions are approximately unbiased estimators of the corresponding measures when the sample size is large.

This paper proposes the refined approximate unbiased estimators of the measures which are obtained by using the second-order term in the Taylor series expansion.

The improved estimators approach to the true measures faster than the original estimators as the sample size becomes larger.

These are shown by the simulation studies.

AMS Subject Classification: 62H17
Key Words: estimation, marginal homogeneity, square contingency table, Taylor series expansion, unbiased estimators

1. Introduction

Consider an $R \times R$ square contingency table with same row and column clas-

Received: October 30, 2013

§Correspondence author

sifications. Let p_{ij} denote the probability that an observation will fall in the ith row and jth column of the table ($i = 1, 2, \ldots, R; j = 1, 2, \ldots, R$), and let X_1 and X_2 denote the row and column variables respectively. The marginal homogeneity (MH) model is defined by

$$p_{i\cdot} = p_{\cdot i} \quad (i = 1, 2, \ldots, R),$$

where $p_{i\cdot} = \sum_{k=1}^{R} p_{ik}$ and $p_{\cdot i} = \sum_{k=1}^{R} p_{ki}$ (Stuart [4]). This model can be expressed as

$$p^c_{i\cdot} = p^c_{\cdot i} \quad (i = 1, 2, \ldots, R),$$

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

Let for $i = 1, 2, \ldots, R-1$

$$G_{1(i)} = \sum_{s=1}^{i} \sum_{t=i+1}^{R} p_{st} \quad \text{and} \quad G_{2(i)} = \sum_{s=i+1}^{R} \sum_{t=1}^{i} p_{st}.$$

Then, the MH model may further be expressed as

$$G_{1(i)} = G_{2(i)} \quad (i = 1, 2, \ldots, R-1).$$

When the MH model does not hold for the given data, one may be interested in measuring the degree of departure from MH. Tomizawa [6] and Tomizawa and Makii [7] considered the measures $\Psi^{(\lambda)}$ and $\Phi^{(\lambda)}$ to represent the degree of departure from MH for the nominal data, which are expressed by using the power-divergence (Read and Cressie [3], p.15). Assuming that $\{p_{i\cdot} + p_{\cdot i}\}$ are all positive, $\Psi^{(\lambda)}$ is defined by

$$\Psi^{(\lambda)} = \frac{1}{2(2^\lambda - 1)} \sum_{i=1}^{R} \left[p_{i\cdot} \left(\left(\frac{p_{i\cdot}}{\pi^*_i} \right)^\lambda - 1 \right) + p_{\cdot i} \left(\left(\frac{p_{\cdot i}}{\pi^*_i} \right)^\lambda - 1 \right) \right],$$

for $\lambda > -1$, where $\pi^*_i = (p_{i\cdot} + p_{\cdot i})/2$ and $\Psi^{(0)} = \lim_{\lambda \to 0} \Psi^{(\lambda)}$. Also, $\Phi^{(\lambda)}$ is given as follows: assuming that $\{p^c_{i\cdot} + p^c_{\cdot i}\}$ are all positive,

$$\Phi^{(\lambda)} = \frac{1}{2(2^\lambda - 1)} \sum_{i=1}^{R} \left[p^c_{i\cdot} \left(\left(\frac{p^c_{i\cdot}}{\pi^{c*}_i} \right)^\lambda - 1 \right) + p^c_{\cdot i} \left(\left(\frac{p^c_{\cdot i}}{\pi^{c*}_i} \right)^\lambda - 1 \right) \right],$$

for $\lambda > -1$, where $\pi^{c*}_i = (p^c_{i\cdot} + p^c_{\cdot i})/2$ and $\Phi^{(0)} = \lim_{\lambda \to 0} \Phi^{(\lambda)}$. $\Psi^{(\lambda)}$ ($\Phi^{(\lambda)}$) is useful for seeing how far the unconditional (conditional) marginal distributions are distant from those with an MH structure.

Tomizawa, Miyamoto and Ashihara [8] considered the measure $\Gamma^{(\lambda)}$ to represent the degree of departure from MH for the ordinal data. Assuming that $\{G_{1(i)} + G_{2(i)}\}$ are all positive, $\Gamma^{(\lambda)}$ is defined by

$$\Gamma^{(\lambda)} = \frac{1}{2^{\lambda}-1} \sum_{i=1}^{R-1} \left[G_{1(i)}^{*} \left(\left(\frac{G_{1(i)}^{*}}{Q_i^{*}} \right)^{\lambda} - 1 \right) + G_{2(i)}^{*} \left(\left(\frac{G_{2(i)}^{*}}{Q_i^{*}} \right)^{\lambda} - 1 \right) \right],$$

for $\lambda > -1$, where $\Delta = \sum_{i=1}^{R-1} \left(G_{1(i)} + G_{2(i)} \right)$,

$$G_{1(i)}^{*} = \frac{G_{1(i)}}{\Delta}, \quad G_{2(i)}^{*} = \frac{G_{2(i)}}{\Delta}, \quad Q_i^{*} = \frac{1}{2} \left(G_{1(i)}^{*} + G_{2(i)}^{*} \right),$$

and $\Gamma^{(0)} = \lim_{\lambda \to 0} \Gamma^{(\lambda)}$.

The measures ($\Psi^{(\lambda)}$, $\Phi^{(\lambda)}$ and $\Gamma^{(\lambda)}$) must lie between 0 and 1 and the degree of departure from MH increases as the value of measure increases. Also, these measures are useful for comparing between several tables.

Using the first-order term in the Taylor series expansion, the estimated measures with the cell probabilities replaced by the corresponding sample proportions are approximately unbiased estimators of the corresponding measures when the sample size is large. Using the second-order Taylor expansion, Tomizawa, Miyamoto and Ohba [9] improved the approximate unbiased estimators of measures of asymmetry for square contingency tables and Tahata, Tomisato and Tomizawa [5] also gave the improved approximate unbiased estimator of log-odds ratio. So we are now interested in, when the sample size n is not so large, proposing the refined approximate unbiased estimators of $\Psi^{(\lambda)}$, $\Phi^{(\lambda)}$ and $\Gamma^{(\lambda)}$.

The purpose of this paper is to propose the improved approximate unbiased estimators of $\Psi^{(\lambda)}$, $\Phi^{(\lambda)}$ and $\Gamma^{(\lambda)}$. Section 2 gives these estimators. Section 3 shows that the proposed estimators work well in many cases by the simulation studies.

2. Refined Approximate Unbiased Estimators

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

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

where "t" means transpose. Also let $\hat{p}_{ij}$ be sample proportion (i.e., $\hat{p}_{ij} = n_{ij}/n$ where $n = \sum\sum n_{ij}$) 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 ([1], p.589) and Bishop, Fienberg and Holland ([2], 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 $\Psi^{(\lambda)}$ ($\Phi^{(\lambda)}$ and $\Gamma^{(\lambda)}$), i.e., $\hat{\Psi}^{(\lambda)}$ ($\hat{\Phi}^{(\lambda)}$ and $\hat{\Gamma}^{(\lambda)}$) is given by $\Psi^{(\lambda)}$ ($\Phi^{(\lambda)}$ and $\Gamma^{(\lambda)}$) with p_{ij} replaced by $\hat{p}_{ij}$, is asymptotically unbiased estimator of $\Psi^{(\lambda)}$ ($\Phi^{(\lambda)}$ and $\Gamma^{(\lambda)}$) 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., $E(g(\hat{p}))$, is approximately equal to

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

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

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

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

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

where $[\partial^2 g(\hat{p})/\partial \hat{p} \partial \hat{p}^t]$ 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 $\hat{p}_{ij}$.

(1) We now propose the refined estimator of the true measure $\Psi^{(\lambda)}$ as follows:

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

where $[\partial^2 \hat{\Psi}^{(\lambda)} / \partial \hat{p} \partial \hat{p}^t]$ is given by $[\partial^2 \Psi^{(\lambda)} / \partial p \partial p^t]$ with p_{ij} replaced by $\hat{p}_{ij}$. The elements of $[\partial^2 \Psi^{(\lambda)} / \partial p \partial p^t]$ are given as follows:

$$\begin{aligned} \frac{\partial^2 \Psi^{(\lambda)}}{\partial p_{ij} \partial p_{kl}} = \frac{p_{2(i)}}{p_{i\cdot} + p_{\cdot i}} d_{(i,k)} K_{2(i)}^{(\lambda)} - \frac{p_{1(i)}}{p_{i\cdot} + p_{\cdot i}} d_{(i,l)} K_{2(i)}^{(\lambda)} \\ - \frac{p_{2(j)}}{p_{j\cdot} + p_{\cdot j}} d_{(j,k)} K_{1(j)}^{(\lambda)} + \frac{p_{1(j)}}{p_{j\cdot} + p_{\cdot j}} d_{(j,l)} K_{1(j)}^{(\lambda)}, \end{aligned}$$

where $p_{1(i)} = p_{i\cdot}/(p_{i\cdot} + p_{\cdot i})$, $p_{2(i)} = p_{\cdot i}/(p_{i\cdot} + p_{\cdot i})$ and

$$\begin{aligned} K_{1(j)}^{(\lambda)} &= \frac{2^\lambda \lambda(\lambda+1)}{2(2^\lambda - 1)} \left[p_{1(j)} \left((p_{1(j)})^{\lambda-1} + (p_{2(j)})^{\lambda-1} \right) \right], \\ K_{2(i)}^{(\lambda)} &= \frac{2^\lambda \lambda(\lambda+1)}{2(2^\lambda - 1)} \left[p_{2(i)} \left((p_{1(i)})^{\lambda-1} + (p_{2(i)})^{\lambda-1} \right) \right], \end{aligned}$$

$$d_{(a,b)} = \begin{cases} 1 & (a = b), \\ 0 & (a \neq b), \end{cases} \qquad K_{1(j)}^{(0)} = \frac{1}{2(\log 2) p_{2(j)}}, \quad K_{2(i)}^{(0)} = \frac{1}{2(\log 2) p_{1(i)}}.$$

(2) In the similar manner, we shall improve the estimated measure $\hat{\Phi}^{(\lambda)}$ as follows:

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

where $[\partial^2 \hat{\Phi}^{(\lambda)} / \partial \hat{p} \partial \hat{p}^t]$ is given by $[\partial^2 \Phi^{(\lambda)} / \partial p \partial p^t]$ with p_{ij} replaced by $\hat{p}_{ij}$. The elements of $[\partial^2 \Phi^{(\lambda)} / \partial p \partial p^t]$ are given as follows:

$$\begin{aligned} \frac{\partial^2 \Phi^{(\lambda)}}{\partial p_{ij} \partial p_{kl}} = \frac{1}{\delta^2} \Bigg[- \left(K_{ij}^{(\lambda)} + K_{kl}^{(\lambda)} \right) + \Bigg(\frac{p^c_{2(i)}}{p^c_{i\cdot} + p^c_{\cdot i}} d_{(i,k)} L_{1(i)}^{(\lambda)} \\ - \frac{p^c_{1(i)}}{p^c_{i\cdot} + p^c_{\cdot i}} d_{(i,l)} L_{1(i)}^{(\lambda)} - \frac{p^c_{2(j)}}{p^c_{j\cdot} + p^c_{\cdot j}} d_{(j,k)} L_{2(j)}^{(\lambda)} + \frac{p^c_{1(j)}}{p^c_{j\cdot} + p^c_{\cdot j}} d_{(j,l)} L_{2(j)}^{(\lambda)} \Bigg) \Bigg], \end{aligned}$$

for $i \neq j$ and $k \neq l$, and $\partial^2 \Phi^{(\lambda)} / \partial p_{ij} \partial p_{kl} = 0$ for $i = j$ and/or $k = l$, where $p^c_{1(i)} = p^c_{i\cdot}/(p^c_{i\cdot} + p^c_{\cdot i})$, $p^c_{2(i)} = p^c_{\cdot i}/(p^c_{i\cdot} + p^c_{\cdot i})$ and

$$K_{ij}^{(\lambda)} = \frac{2^{\lambda-1}}{2^\lambda - 1}\Bigg[(p_{1(i)}^c)^\lambda + (p_{2(j)}^c)^\lambda + \lambda p_{2(i)}^c \left((p_{1(i)}^c)^\lambda - (p_{2(i)}^c)^\lambda\right) - \lambda p_{1(j)}^c \left((p_{1(j)}^c)^\lambda - (p_{2(j)}^c)^\lambda\right) - \frac{2\left((2^\lambda - 1)\Phi^{(\lambda)} + 1\right)}{2^\lambda}\Bigg],$$

$$L_{1(i)}^{(\lambda)} = \frac{\lambda 2^{\lambda-1}}{2^\lambda - 1}\left[(p_{1(i)}^c)^{\lambda-1} + \lambda p_{2(i)}^c \left((p_{1(i)}^c)^{\lambda-1} + (p_{2(i)}^c)^{\lambda-1}\right) - \left((p_{1(i)}^c)^\lambda - (p_{2(i)}^c)^\lambda\right)\right],$$

$$L_{2(i)}^{(\lambda)} = \frac{\lambda 2^{\lambda-1}}{2^\lambda - 1}\left[(p_{2(i)}^c)^{\lambda-1} + \lambda p_{1(i)}^c \left((p_{1(i)}^c)^{\lambda-1} + (p_{2(i)}^c)^{\lambda-1}\right) - \left((p_{2(i)}^c)^\lambda - (p_{1(i)}^c)^\lambda\right)\right],$$

$$K_{ij}^{(0)} = \frac{\log(p_{1(i)}^c p_{2(j)}^c)}{2\log 2} + (1 - \Phi^{(0)}),$$

$$L_{1(i)}^{(0)} = \frac{1}{2(\log 2)p_{1(i)}^c}, \quad L_{2(i)}^{(0)} = \frac{1}{2(\log 2)p_{2(i)}^c}.$$

(3) Finally, we shall improve the estimated measure $\hat{\Gamma}^{(\lambda)}$ as follows:

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

where $[\partial^2 \hat{\Gamma}^{(\lambda)}/\partial \hat{p}\partial \hat{p}^t]$ is given by $[\partial^2 \Gamma^{(\lambda)}/\partial p \partial p^t]$ with p_{ij} replaced by $\hat{p}_{ij}$. The elements of $[\partial^2 \Gamma^{(\lambda)}/\partial p \partial p^t]$ are given as follows:

$$\frac{\partial^2 \Gamma^{(\lambda)}}{\partial p_{ij} \partial p_{kl}} = \frac{1}{\Delta}\left[I_{(i<j)}\left(L_{1(ij;kl)}^{(\lambda)} - |k-l| W_{1(ij)}^{(\lambda)}\right) + I_{(i>j)}\left(L_{2(ij;kl)}^{(\lambda)} - |k-l| W_{2(ij)}^{(\lambda)}\right)\right],$$

where $I_{(\cdot)} = 1$ if true, 0 if not, and

$$W_{1(ij)}^{(\lambda)} = \frac{2^\lambda}{\Delta(2^\lambda - 1)}\left[\sum_{s=i}^{j-1}\left\{(G_{1(s)}^c)^\lambda + \lambda\left((G_{1(s)}^c)^\lambda - (G_{2(s)}^c)^\lambda\right) G_{2(s)}^c\right\}\right.$$

Table 1: (a) The artificial probabilities $\{p_{ij}\}$, (b) the values of $\Psi^{(\lambda)}$, $\Phi^{(\lambda)}$ and $\Gamma^{(\lambda)}$, and (c) the each sample mean of the values of estimated measures obtained by generating 10000 times simulations, with each sample size n, for Table 1(a)

(a)

	(1)	(2)	(3)	(4)
(1)	0.050	0.043	0.036	0.037
(2)	0.105	0.050	0.122	0.072
(3)	0.042	0.024	0.050	0.030
(4)	0.118	0.051	0.120	0.050

(b)

λ	$\Psi^{(\lambda)}$	$\Phi^{(\lambda)}$	$\Gamma^{(\lambda)}$
1.0	0.1110	0.1737	0.0861
1.5	0.1136	0.1775	0.0880

(c)

n	λ	$\hat{\Psi}^{(\lambda)}$	$\hat{\Psi}^{*(\lambda)}$	$\hat{\Phi}^{(\lambda)}$	$\hat{\Phi}^{*(\lambda)}$	$\hat{\Gamma}^{(\lambda)}$	$\hat{\Gamma}^{*(\lambda)}$
50	1.0	0.1399	0.1124	0.2148	0.1754	0.1270	0.0869
	1.5	0.1418	0.1140	0.2188	0.1795	0.1301	0.0894
100	1.0	0.1260	0.1120	0.1937	0.1735	0.1070	0.0865
	1.5	0.1275	0.1134	0.1975	0.1773	0.1086	0.0878
200	1.0	0.1181	0.1110	0.1839	0.1737	0.0963	0.0860
	1.5	0.1210	0.1139	0.1877	0.1775	0.0988	0.0884
400	1.0	0.1147	0.1111	0.1791	0.1739	0.0913	0.0861
	1.5	0.1171	0.1135	0.1828	0.1777	0.0937	0.0885
800	1.0	0.1129	0.1111	0.1765	0.1739	0.0889	0.0863
	1.5	0.1154	0.1136	0.1803	0.1777	0.0905	0.0878

$$\left. -(j-i)\frac{(2^{\lambda}-1)\Gamma^{(\lambda)}+1}{2^{\lambda}}\right],$$

$$W_{2(ij)}^{(\lambda)} = \frac{2^{\lambda}}{\Delta(2^{\lambda}-1)}\left[\sum_{s=j}^{i-1}\left\{(G_{2(s)}^{c})^{\lambda}+\lambda\left((G_{2(s)}^{c})^{\lambda}-(G_{1(s)}^{c})^{\lambda}\right)G_{1(s)}^{c}\right\}\right.$$

$$\left. -(i-j)\frac{(2^{\lambda}-1)\Gamma^{(\lambda)}+1}{2^{\lambda}}\right],$$

Table 2: (a) The artificial probabilities $\{p_{ij}\}$, (b) the values of $\Psi^{(\lambda)}$, $\Phi^{(\lambda)}$ and $\Gamma^{(\lambda)}$, and (c) the each sample mean of the values of estimated measures obtained by generating 10000 times simulations, with each sample size n, for Table 2(a)

(a)

	(1)	(2)	(3)	(4)
(1)	0.003	0.037	0.229	0.145
(2)	0.023	0.003	0.141	0.245
(3)	0.021	0.019	0.003	0.043
(4)	0.025	0.031	0.029	0.003

(b)

λ	$\Psi^{(\lambda)}$	$\Phi^{(\lambda)}$	$\Gamma^{(\lambda)}$
1.0	0.4415	0.4523	0.5353
1.5	0.4483	0.4591	0.5423

(c)

n	λ	$\hat{\Psi}^{(\lambda)}$	$\hat{\Psi}^{*(\lambda)}$	$\hat{\Phi}^{(\lambda)}$	$\hat{\Phi}^{*(\lambda)}$	$\hat{\Gamma}^{(\lambda)}$	$\hat{\Gamma}^{*(\lambda)}$
50	1.0	0.4649	0.4438	0.4759	0.4547	0.5494	0.5347
	1.5	0.4691	0.4488	0.4821	0.4617	0.5561	0.5422
100	1.0	0.4527	0.4419	0.4626	0.4517	0.5450	0.5376
	1.5	0.4577	0.4473	0.4689	0.4585	0.5498	0.5427
200	1.0	0.4479	0.4424	0.4580	0.4525	0.5392	0.5355
	1.5	0.4524	0.4472	0.4641	0.4588	0.5459	0.5424
400	1.0	0.4441	0.4414	0.4551	0.4524	0.5371	0.5353
	1.5	0.4504	0.4478	0.4612	0.4585	0.5446	0.5428
800	1.0	0.4429	0.4415	0.4535	0.4522	0.5361	0.5351
	1.5	0.4499	0.4486	0.4602	0.4588	0.5428	0.5419

$$
\begin{aligned}
L_{1(ij;kl)}^{(\lambda)} = \frac{2^\lambda}{2^\lambda - 1} \sum_{s=i}^{j-1} \Big[& \frac{d_{kl(s)} G_{2(s)}^c - d_{lk(s)} G_{1(s)}^c}{G_{1(s)} + G_{2(s)}} \lambda \Big((G_{1(s)}^c)^{\lambda-1} \\
& + \lambda G_{2(s)}^c \Big((G_{1(s)}^c)^{\lambda-1} + (G_{2(s)}^c)^{\lambda-1} \Big) - ((G_{1(s)}^c)^\lambda - (G_{2(s)}^c)^\lambda) \Big) \Big] \\
& - (j-i)(I_{(k<l)} W_{1(kl)}^{(\lambda)} + I_{(k>l)} W_{2(kl)}^{(\lambda)}),
\end{aligned}
$$

$$
\begin{aligned}
L_{2(ij;kl)}^{(\lambda)} = \frac{2^\lambda}{2^\lambda - 1} \sum_{s=j}^{i-1} \Big[& \frac{d_{kl(s)} G_{2(s)}^c - d_{lk(s)} G_{1(s)}^c}{G_{1(s)} + G_{2(s)}} \lambda \Big(-(G_{2(s)}^c)^{\lambda-1} \\
& - \lambda G_{1(s)}^c \Big((G_{1(s)}^c)^{\lambda-1} + (G_{2(s)}^c)^{\lambda-1} \Big) + ((G_{2(s)}^c)^\lambda - (G_{1(s)}^c)^\lambda) \Big) \Big]
\end{aligned}
$$

Table 3: (a) The artificial probabilities $\{p_{ij}\}$, (b) the values of $\Psi^{(\lambda)}$, $\Phi^{(\lambda)}$ and $\Gamma^{(\lambda)}$, and (c) the each sample mean of the values of estimated measures obtained by generating 10000 times simulations, with each sample size n, for Table 3(a)

(a)

	(1)	(2)	(3)	(4)
(1)	0.002	0.009	0.226	0.219
(2)	0.008	0.002	0.251	0.224
(3)	0.007	0.008	0.002	0.022
(4)	0.005	0.006	0.007	0.002

(b)

λ	$\Psi^{(\lambda)}$	$\Phi^{(\lambda)}$	$\Gamma^{(\lambda)}$
1.0	0.7999	0.8129	0.8702
1.5	0.8047	0.8174	0.8737

(c)

n	λ	$\hat{\Psi}^{(\lambda)}$	$\hat{\Psi}^{*(\lambda)}$	$\hat{\Phi}^{(\lambda)}$	$\hat{\Phi}^{*(\lambda)}$	$\hat{\Gamma}^{(\lambda)}$	$\hat{\Gamma}^{*(\lambda)}$
50	1.0	0.8056	0.7980	0.8180	0.8108	0.8747	0.8709
	1.5	0.8110	0.8042	0.8232	0.8167	0.8762	0.8728
100	1.0	0.8047	0.8009	0.8164	0.8128	0.8720	0.8700
	1.5	0.8082	0.8047	0.8181	0.8174	0.8763	0.8747
200	1.0	0.8020	0.8001	0.8152	0.8134	0.8718	0.8708
	1.5	0.8065	0.8048	0.8189	0.8172	0.8743	0.8735
400	1.0	0.8011	0.8001	0.8139	0.8130	0.8708	0.8703
	1.5	0.8054	0.8045	0.8179	0.8171	0.8741	0.8737
800	1.0	0.8002	0.7997	0.8132	0.8128	0.8707	0.8704
	1.5	0.8054	0.8050	0.8180	0.8176	0.8742	0.8740

$$- (i-j)\left(I_{(k<l)} W_{1(kl)}^{(\lambda)} + I_{(k>l)} W_{2(kl)}^{(\lambda)}\right),$$

with $G_{1(i)}^c = G_{1(i)}/(G_{1(i)} + G_{2(i)})$, $G_{2(i)}^c = G_{2(i)}/(G_{1(i)} + G_{2(i)})$,

$$d_{kl(s)} = \begin{cases} 1 & (k \le s, l \ge s+1), \\ 0 & (\text{otherwise}), \end{cases}$$

and

$$W_{1(ij)}^{(0)} = \frac{1}{\Delta \log 2}\left[\sum_{s=i}^{j-1} \log G_{1(s)}^c - (j-i)(\Gamma^{(0)} - 1)\log 2\right],$$

$$W_{2(ij)}^{(0)} = \frac{1}{\Delta \log 2}\left[\sum_{s=j}^{i-1} \log G_{2(s)}^c - (i-j)(\Gamma^{(0)} - 1)\log 2\right],$$

Table 4: (a) The artificial probabilities $\{p_{ij}\}$, (b) the values of $\Psi^{(\lambda)}$, $\Phi^{(\lambda)}$ and $\Gamma^{(\lambda)}$, and (c) the each sample mean of the values of estimated measures obtained by generating 10000 times simulations, with each sample size n, for Table 4(a)

(a)

	(1)	(2)	(3)	(4)
(1)	0.100	0.037	0.099	0.052
(2)	0.062	0.100	0.086	0.096
(3)	0.013	0.018	0.100	0.051
(4)	0.018	0.020	0.048	0.100

(b)

λ	$\Psi^{(\lambda)}$	$\Phi^{(\lambda)}$	$\Gamma^{(\lambda)}$
1.0	0.0722	0.1990	0.2586
1.5	0.0739	0.2031	0.2632

(c)

n	λ	$\hat{\Psi}^{(\lambda)}$	$\hat{\Psi}^{*(\lambda)}$	$\hat{\Phi}^{(\lambda)}$	$\hat{\Phi}^{*(\lambda)}$	$\hat{\Gamma}^{(\lambda)}$	$\hat{\Gamma}^{*(\lambda)}$
50	1.0	0.0953	0.0735	0.2547	0.2043	0.3080	0.2642
	1.5	0.0974	0.0754	0.2602	0.2100	0.3134	0.2699
100	1.0	0.0833	0.0723	0.2256	0.1995	0.2812	0.2586
	1.5	0.0858	0.0746	0.2306	0.2046	0.2867	0.2644
200	1.0	0.0778	0.0722	0.2123	0.1991	0.2698	0.2584
	1.5	0.0801	0.0745	0.2182	0.2050	0.2763	0.2650
400	1.0	0.0753	0.0725	0.2063	0.1997	0.2649	0.2592
	1.5	0.0764	0.0736	0.2089	0.2022	0.2679	0.2622
800	1.0	0.0738	0.0724	0.2028	0.1995	0.2624	0.2595
	1.5	0.0752	0.0738	0.2064	0.2031	0.2658	0.2629

$$L^{(0)}_{1(ij;kl)} = \frac{1}{\log 2}\left[\sum_{s=i}^{j-1}\frac{d_{kl(s)}G^c_{2(s)} - d_{lk(s)}G^c_{1(s)}}{G^c_{1(s)}(G_{1(s)}+G_{2(s)})}\right] - (j-i)(I_{(k<l)}W^{(0)}_{1(kl)} + I_{(k>l)}W^{(0)}_{2(kl)}),$$

$$L^{(0)}_{2(ij;kl)} = \frac{-1}{\log 2}\left[\sum_{s=j}^{i-1}\frac{d_{kl(s)}G^c_{2(s)} - d_{lk(s)}G^c_{1(s)}}{G^c_{2(s)}(G_{1(s)}+G_{2(s)})}\right] - (i-j)(I_{(k<l)}W^{(0)}_{1(kl)} + I_{(k>l)}W^{(0)}_{2(kl)}).$$

Table 5: (a) The artificial probabilities $\{p_{ij}\}$, (b) the values of $\Psi^{(\lambda)}$, $\Phi^{(\lambda)}$ and $\Gamma^{(\lambda)}$, and (c) the each sample mean of the values of estimated measures obtained by generating 10000 times simulations, with each sample size n, for Table 5(a)

(a)

	(1)	(2)	(3)	(4)
(1)	0.200	0.017	0.055	0.029
(2)	0.010	0.200	0.038	0.059
(3)	0.010	0.014	0.200	0.019
(4)	0.025	0.011	0.013	0.100

(b)

λ	$\Psi^{(\lambda)}$	$\Phi^{(\lambda)}$	$\Gamma^{(\lambda)}$
1.0	0.0151	0.1634	0.1912
1.5	0.0154	0.1670	0.1952

(c)

n	λ	$\hat{\Psi}^{(\lambda)}$	$\hat{\Psi}^{*(\lambda)}$	$\hat{\Phi}^{(\lambda)}$	$\hat{\Phi}^{*(\lambda)}$	$\hat{\Gamma}^{(\lambda)}$	$\hat{\Gamma}^{*(\lambda)}$
50	1.0	0.0276	0.0153	0.2788	0.1804	0.2920	0.2081
	1.5	0.0285	0.0159	0.2838	0.1854	0.2935	0.2097
100	1.0	0.0214	0.0152	0.2189	0.1659	0.2386	0.1938
	1.5	0.0217	0.0155	0.2246	0.1715	0.2443	0.1995
200	1.0	0.0182	0.0151	0.1901	0.1628	0.2155	0.1926
	1.5	0.0186	0.0155	0.1955	0.1682	0.2188	0.1959
400	1.0	0.0167	0.0151	0.1779	0.1641	0.2037	0.1922
	1.5	0.0171	0.0155	0.1806	0.1668	0.2066	0.1950
800	1.0	0.0159	0.0151	0.1702	0.1633	0.1963	0.1904
	1.5	0.0163	0.0155	0.1739	0.1669	0.2008	0.1950

3. Simulation Studies

By the simulation studies, we calculate the values of estimated measures $\hat{\Psi}^{(\lambda)}$, $\hat{\Psi}^{*(\lambda)}$, $\hat{\Phi}^{(\lambda)}$, $\hat{\Phi}^{*(\lambda)}$, $\hat{\Gamma}^{(\lambda)}$ and $\hat{\Gamma}^{*(\lambda)}$ from the observed frequencies of sample size $n = 50$, 100, 200, 400 and 800, which are obtained from the true probability distribution. We shall compare the each sample mean of the values of $\hat{\Psi}^{(\lambda)}$ and $\hat{\Psi}^{*(\lambda)}$ obtained by 10000 times simulations, for each sample size. Similarly, the sample mean of the value of $\hat{\Phi}^{(\lambda)}$ ($\hat{\Gamma}^{(\lambda)}$) is compared to that of $\hat{\Phi}^{*(\lambda)}$ ($\hat{\Gamma}^{*(\lambda)}$). The results of simulations are given in Tables 1 to 5.

Tables 1(a), 2(a) and 3(a) have a characteristic that the sum of the probabilities of main-diagonal cells is very small. Also the true values of measures for Tables 1(a), 2(a) and 3(a) are small, medium and large, respectively. The sums of the probabilities of main-diagonal cells for Tables 4(a) and 5(a) are greater than those for Tables 1(a), 2(a) and 3(a), and the true value of measures for Tables 4(a) and 5(a) are small.

We can see from Tables 1(c) to 5(c) that the refined estimator $\hat{\Psi}^{*(\lambda)}$ approaches to the true value $\Psi^{(\lambda)}$ faster than the former estimator $\hat{\Psi}^{(\lambda)}$ when $\lambda \geq 1$. Also we can obtain the similar results for the improved estimators $\hat{\Phi}^{*(\lambda)}$ and $\hat{\Gamma}^{*(\lambda)}$.

4. Concluding Remarks

This paper has proposed the refined approximate unbiased estimators $\hat{\Psi}^{*(\lambda)}$, $\hat{\Phi}^{*(\lambda)}$ and $\hat{\Gamma}^{*(\lambda)}$ of the true measures $\Psi^{(\lambda)}$, $\Phi^{(\lambda)}$ and $\Gamma^{(\lambda)}$, respectively.

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

When $\lambda < 1$, we can calculate the refined estimators $\hat{\Psi}^{*(\lambda)}$, $\hat{\Phi}^{*(\lambda)}$ and $\hat{\Gamma}^{*(\lambda)}$ only the case of $\{\hat{p}_{i\cdot} \neq 0, \hat{p}_{\cdot i} \neq 0\}$, $\{\hat{p}^c_{i\cdot} \neq 0, \hat{p}^c_{\cdot i} \neq 0\}$ and $\{\hat{G}_{1(i)} \neq 0, \hat{G}_{2(i)} \neq 0\}$, respectively. On the other hand, the former estimators $\hat{\Psi}^{(\lambda)}$, $\hat{\Phi}^{(\lambda)}$ and $\hat{\Gamma}^{(\lambda)}$ can be calculated for the case of $\{\hat{p}_{i\cdot} + \hat{p}_{\cdot i} \neq 0\}$, $\{\hat{p}^c_{i\cdot} + \hat{p}^c_{\cdot i} \neq 0\}$ and $\{\hat{G}_{1(i)} + \hat{G}_{2(i)} \neq 0\}$, respectively. This means that the calculable conditions of the refined estimators are different from that of the former estimators. Thus, it seems difficult to evaluate whether the improved estimators tend to approach to the true value faster than the former estimators by simulation study when $\lambda < 1$.

Therefore, we recommend that the proposed estimators should be used for the case of $\lambda \geq 1$. Then these estimators work very well.

Acknowledgments

The authors would like to thank two anonymous referees for helpful comments.

References

[1] A. Agresti, *Categorical Data Analysis*, 3rd edition, Wiley, Hoboken, New Jersey (2013).

[2] Y.M.M. Bishop, S.E. Fienberg and P.W. Holland, *Discrete Multivariate Analysis: Theory and Practice*, The MIT Press, Cambridge, Massachusetts (1975).

[3] T.R.C. Read and N. Cressie, *Goodness-of-Fit Statistics for Discrete Multivariate Data*, Springer, New York (1988).

[4] A. Stuart, A test for homogeneity of the marginal distributions in a two-way classification, *Biometrika*, **42** (1955), 412-416.

[5] K. Tahata, R. Tomisato and S. Tomizawa, An improved approximate unbiased estimator of log-odds ratio for 2×2 contingency tables, *Advances and Applications in Statistics*, **9** (2008), 1-12.

[6] S. Tomizawa, Measures of departure from marginal homogeneity for contingency tables with nominal categories, *Journal of the Royal Statistical Society, Series D; The Statistician*, **44** (1995), 425-439.

[7] S. Tomizawa and T. Makii, Generalized measures of departure from marginal homogeneity for contingency tables with nominal categories, *Journal of Statistical Research*, **35** (2001), 1-24.

[8] S. Tomizawa, N. Miyamoto and N. Ashihara, Measure of departure from marginal homogeneity for square contingency tables having ordered categories, *Behaviormetrika*, **30** (2003), 173-193.

[9] S. Tomizawa, N. Miyamoto and N. Ohba, Improved approximate unbiased estimators of measures of asymmetry for square contingency tables, *Advances and Applications in Statistics*, **7** (2007), 47-63.