STATISTICA, anno LXXXI, n. 2, 2021

POLYNOMIAL COLUMNS-PARAMETER SYMMETRY MODEL AND ITS DECOMPOSITION FOR SQUARE CONTINGENCY TABLES

Shuji Ando [1]
Department of Information and Computer Technology, Tokyo University of Science, Tokyo, Japan

1. INTRODUCTION

Consider an $R \times R$ square contingency table with ordered categories. Let p_{ij} denote the probability that an observation will fall in the (i, j)th cell of the table. The symmetry (S) model (Bowker, 1948) is defined by

$$\log\left(\frac{p_{ij}}{p_{ji}}\right) = 0 \quad (i < j).$$

In the S model, the log odds, $\log(p_{ij}/p_{ji})$ for all $i < j$, are zero. The conditional symmetry (CS) model (McCullagh, 1978) is defined by

$$\log\left(\frac{p_{ij}}{p_{ji}}\right) = \delta \quad (i < j),$$

where the parameter δ is unspecified. The CS model with $\delta = 0$ is equivalent to the S model. Thus, the CS model is more parsimonious than the S model. In the CS model, the log odds, $\log(p_{ij}/p_{ji})$ for all $i < j$, are constant (i.e., δ).

The diagonals-parameter symmetry (DPS) model (Goodman, 1979) is defined by

$$\log\left(\frac{p_{ij}}{p_{ji}}\right) = \delta_{j-i} \quad (i < j),$$

where the parameters $\{\delta_{j-i}\}$ are unspecified.

[1] Corresponding Author. E-mail: shuji.ando@rs.tus.ac.jp

The DPS models with $\delta_1 = \delta_2 = \cdots = \delta_{R-1} = 0$ and $\delta_1 = \delta_2 = \cdots = \delta_{R-1} = \delta$ are equivalent to the S and CS models, respectively. Moreover, the DPS model with $\delta_{j-i} = (j-i)\delta$, for all $i < j$, is equivalent to the linear diagonals-parameter symmetry (LDPS) model (Agresti, 1983). Thus, the CS and LDPS models are more parsimonious than the DPS model. In the DPS and LDPS models, the log odds, $\log(p_{ij}/p_{ji})$ for all $i < j$, depend on the difference between j and i, although they are zero under the S model, and constant under the CS model.

As a generalized model including these models, Tomizawa (1990b) proposed the polynomial diagonals-parameter symmetry (PDPS) model, defined by

$$\log\left(\frac{p_{ij}}{p_{ji}}\right) = \sum_{k=1}^{R-1} (j-i)^{k-1} \Delta_k \quad (i < j),$$

where the parameters $\{\Delta_k\}$ are unspecified. In the PDPS model, the log odds, $\log(p_{ij}/p_{ji})$ for all $i < j$, are a polynomial of degree of $R-2$ with respect to the difference between j and i. The PDPS model can represent more parsimonious models than the DPS model by setting some parameters $\{\Delta_k\}$ to zero. For example, the PDPS models that are set $\Delta_1 = \Delta_2 = \cdots = \Delta_{R-1} = 0$, $\Delta_2 = \Delta_3 = \cdots = \Delta_{R-1} = 0$ and $\Delta_1 = \Delta_3 = \cdots = \Delta_{R-1} = 0$ are equivalent to the S, CS, and LDPS models, respectively. It must be noted that the PDPS model with $\Delta_k \neq 0$, for all $k = 1, 2, \ldots, R-1$, is equivalent to the DPS model (see, e.g., Tomizawa, 1990b, 1991; Tahata and Tomizawa, 2014).

In previous studies, many decomposition theorems of model were given, and Tahata and Tomizawa (2014) summarized over 30 the decomposition theorems. The decomposition theorem is useful for explaining why the model does not hold, it is one of the important disciplines of the research on square contingency tables. In order to provide the decomposition theorem of the PDPS model using the DPS model, Tomizawa (1991) introduced the polynomial diagonals-marginal symmetry (PDMS) model, defined by

$$\log\left(\frac{p_d^+}{p_d^-}\right) = \sum_{k=1}^{R-1} d^{k-1} \Delta_k \quad (d = 1, 2, \ldots, R-1),$$

where

$$p_d^+ = \sum_{l=1}^{R-d} p_{l,l+d} \quad \text{and} \quad p_d^- = \sum_{l=1}^{R-d} p_{l+d,l}.$$

It must be noted that the parameters $\{\Delta_k\}$ of the PDMS model are unspecified. The PDMS model can represent various models by setting some parameters $\{\Delta_k\}$ to zero as well as the PDPS model. For example, the PDMS models that are set $\Delta_2 = \Delta_3 = \cdots = \Delta_{R-1} = 0$ and $\Delta_1 = \Delta_3 = \cdots = \Delta_{R-1} = 0$ are equivalent to the conditional diagonals-marginal symmetry (CDMS) (Tomizawa, 1987) and linear diagonals-marginal symmetry (LDMS) (Tomizawa, 1990a) models, respectively.

Tomizawa (1991) showed that (i) it is necessary to hold the PDMS model, in addition to the DPS model, to satisfy the PDPS model, and (ii) the value of likelihood ratio chi-square statistic for testing the PDPS model is equal to the sum of that for testing the DPS and PDMS models.

Let f_{ij} denote the observed frequency in the (i, j)th cell of the table. We consider the data set in Table 1 obtained from Hashimoto (2018, pp. 122–123), which presents the cross-classification of occupational status categories for father and son dyads in Japan examined in 2015. Table 2 shows the values of f_{ij}/f_{ji} for all $i < j$ in Table 1. From Table 2, it is likely that the values of f_{ij}/f_{ji} for all $i < j$ will depend on only the value of column category j (rather than the difference between j and i).

TABLE 1
Cross-classification of occupational status categories for father and son dyads in Japan, obtained from Hashimoto (2018, pp. 122–123).

	Son status				
Father status	(1)	(2)	(3)	(4)	Total
Highest (1)	33	31	10	7	81
(2)	6	151	81	25	263
(3)	9	131	207	25	372
Lowest (4)	24	97	95	62	278
Total	72	410	393	119	994

TABLE 2
The values of f_{ij}/f_{ji} for all $i < j$ in Table 1.

f_{ij}/f_{ji}	j			
i	1	2	3	4
1	—	5.167	1.111	0.292
2	—	—	0.618	0.258
3	—	—	—	0.263
4	—	—	—	—

This study proposes a model that the log odds, $\log(p_{ij}/p_{ji})$ for all $i < j$, are a polynomial with respect to the column category j. Moreover, this study (i) provides the decomposition theorem of the proposed model, and (ii) shows the value of likelihood ratio chi-square statistic for testing the proposed model is equal to the sum of that for testing the decomposed two models.

The remainder of this paper is organized as follows. Section 2 proposes a polynomial columns-parameter symmetry model. Section 3 provides a decomposition theorem of the proposed model. Section 4 shows the value of likelihood ratio chi-square statistic for testing the proposed model is equal to the sum of that for testing the decomposed two models. Section 5 demonstrates the utility of the proposed model using real-world data in Table 1. Section 6 closes with concluding remarks.

2. Proposed model

Tomizawa (1985) proposed the columns-parameter symmetry (CPS) model, defined by

$$\log\left(\frac{p_{ij}}{p_{ji}}\right) = \delta_{j-1} \quad (i < j),$$

where the parameters $\{\delta_{j-1}\}$ are unspecified. The CPS models with $\delta_1 = \delta_2 = \cdots = \delta_{R-1} = 0$ and $\delta_1 = \delta_2 = \cdots = \delta_{R-1} = \delta$ are equivalent to the S and CS models, respectively. Moreover, the CPS model with $\delta_{j-1} = (j-1)\delta$, for all $i < j$, is equivalent to the linear columns-parameter symmetry (LCPS) model (Tomizawa *et al.*, 2006). In the CPS and LCPS models, the log odds, $\log(p_{ij}/p_{ji})$ for all $i < j$, depend on only the value of column category j. It must be noted that Tomizawa (1985) referred to the CPS model as the odds-symmetry model.

As a generalized model including the CPS and LCPS models, we propose the polynomial columns-parameter symmetry (PCPS) model, defined by

$$\log\left(\frac{p_{ij}}{p_{ji}}\right) = \sum_{k=1}^{R-1}(j-1)^{k-1}\Delta_k \quad (i < j),$$

where the parameters $\{\Delta_k\}$ are unspecified. In the PCPS model, the log odds, $\log(p_{ij}/p_{ji})$ for all $i < j$, are a polynomial of degree of $R-2$ with respect to the column category j. The PCPS model can represent more parsimonious models than the CPS model by setting some parameters $\{\Delta_k\}$ to zero. For example, the PCPS models that are set $\Delta_1 = \Delta_2 = \cdots = \Delta_{R-1} = 0$, $\Delta_2 = \Delta_3 = \cdots = \Delta_{R-1} = 0$ and $\Delta_1 = \Delta_3 = \cdots = \Delta_{R-1} = 0$ are equivalent to the S, CS, and LCPS models, respectively.

The connection between $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_{R-1})^t$ in the CPS model and $\boldsymbol{\Delta} = (\Delta_1, \Delta_2, \ldots, \Delta_{R-1})^t$ in the PCPS model is given as

$$\boldsymbol{\delta} = V\boldsymbol{\Delta},$$

where

$$V = \begin{pmatrix} 1 & 1 & 1^2 & \cdots & 1^{R-2} \\ 1 & 2 & 2^2 & \cdots & 2^{R-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & R-1 & (R-1)^2 & \cdots & (R-1)^{R-2} \end{pmatrix}.$$

Since V is the Vandermonde matrix of order $(R-1)\times(R-1)$, the transformation from the $\boldsymbol{\delta}$ in the CPS model to the $\boldsymbol{\Delta}$ in the PCPS model is one-to-one. Therefore, the PCPS model with $\Delta_k \neq 0$, for all $k = 1, 2, \ldots, R-1$, is equivalent to the CPS model.

3. Decomposition of proposed model

In this Section, we focus on the relationship between the CPS and PCPS models. Under some parameters $\{\Delta_k\}$ in the PCPS model is set to zero, the CPS model constantly holds when the PCPS model holds. However, the converse is not necessarily true. We want to detect a model that it is necessary to satisfy, in addition to the CPS model, to satisfy the PCPS model. In other words, we want to provide the decomposition theorem of the PCPS model using the CPS model.

In order to provide the decomposition theorem of the PCPS model using the CPS model, we propose the polynomial columns-marginal symmetry (PCMS) model, defined by

$$\log\left(\frac{p_c^+}{p_c^-}\right) = \sum_{k=1}^{R-1} c^{k-1}\Delta_k \quad (c = 1, 2, \dots, R-1),$$

where

$$p_c^+ = \sum_{l=1}^{c-1} p_{lc} \quad \text{and} \quad p_c^- = \sum_{l=1}^{c-1} p_{cl}.$$

It must be noted that the parameters $\{\Delta_k\}$ of the PCMS model are unspecified. The PCMS model can represent various models by setting some parameters $\{\Delta_k\}$ to zero as well as the PCPS model. For example, the PCMS models that are set $\Delta_1 = \Delta_2 = \cdots = \Delta_{R-1} = 0$ and $\Delta_2 = \Delta_3 = \cdots = \Delta_{R-1} = 0$ are equivalent to the columns-marginal symmetry (CMS) and conditional columns-marginal symmetry (CCMS) models (Tomizawa, 1984).

Let $U(M)$ imply that model M holds. We provide the following decomposition theorem of the PCPS model using the CPS and PCMS models.

Theorem 1. *Under some parameters $\{\Delta_k\}$ in the PCPS and PCMS models are set to zero, the following necessary and sufficient condition holds:*

$$U(PCPS) \Leftrightarrow U(CPS) \wedge U(PCMS).$$

Proof. It is clear that the necessary condition $U(PCPS) \Rightarrow U(CPS) \wedge U(PCMS)$ holds. It is necessary to show that the sufficient condition $U(PCPS) \Leftarrow U(CPS) \wedge U(PCMS)$ holds. Since the both CPS and PCMS models hold, the following equality holds:

$$\delta_c = \sum_{k=1}^{R-1} c^{k-1}\Delta_k \quad (c = 1, 2, \dots, R-1).$$

Therefore, we obtain the following equality:

$$\log\left(\frac{p_{ij}}{p_{ji}}\right) = \sum_{k=1}^{R-1} (j-1)^{k-1}\Delta_k \quad (i < j).$$

The proof is completed. □

When all Δ_k (or all Δ_k except Δ_1) in the PCPS and PCMS models are set to zero, Theorem 1 is equivalent to the decomposition theorem of the S (or CS) model proposed by Ando and Aoba (2018).

4. Goodness-of-fit test for proposed model and model selection

4.1. Goodness-of-fit test

Assume that a multinomial distribution applies to the $R \times R$ table. The maximum likelihood estimates (MLEs) of the expected frequencies under the model can be obtained using, for example, the Newton-Raphson method in the log-likelihood equation.

In order to obtain the MLEs of the expected frequencies under the PCPS model, we must maximize the following Lagrangian

$$\sum_{i=1}^{R}\sum_{j=1}^{R} f_{ij}\log p_{ij} - \phi\left(\sum_{i=1}^{R}\sum_{j=1}^{R} p_{ij} - 1\right) - \sum_{i<j}\psi_{ij}\left(\frac{p_{ij}}{p_{ji}} - \exp\left[\sum_{k=1}^{R-1}(j-1)^{k-1}\Delta_k\right]\right),$$

with respect to $\{p_{ij}\}, \phi, \{\psi_{ij}\}$, and $\{\Delta_k\}$. Similarly, the MLEs of the expected frequencies under the PCMS are obtained by maximizing the following Lagrangian

$$\sum_{i=1}^{R}\sum_{j=1}^{R} f_{ij}\log p_{ij} - \phi\left(\sum_{i=1}^{R}\sum_{j=1}^{R} p_{ij} - 1\right) - \sum_{c=1}^{R-1}\psi_{c}\left(\frac{p_c^{+}}{p_c^{-}} - \exp\left[\sum_{k=1}^{R-1}c^{k-1}\Delta_k\right]\right),$$

with respect to $\{p_{ij}\}, \phi, \{\psi_c\}$, and $\{\Delta_k\}$.

Each model can be tested for the goodness-of-fit by, for example, the likelihood ratio chi-square statistic (denoted by G^2) with the corresponding degrees of freedom. The test statistic G^2 of model M is given by

$$G^2(M) = 2\sum_{i=1}^{R}\sum_{j=1}^{R} f_{ij}\log\left(\frac{f_{ij}}{\hat{e}_{ij}}\right),$$

where $\hat{e}_{ij}$ is the MLE of the expected frequency e_{ij} under model M. Let γ be the number of Δ_k set to zero in the PCPS and PCMS models. The number of degrees of freedom for the CPS, PCPS, PCMS models are $(R-1)(R-2)/2$, $(R-1)(R-2)/2+\gamma$, and γ, respectively. It must be noted that the number of degrees of freedom for the CPS model is equal to the sum of that for the PCPS and PCMS models. We provide the following theorem.

THEOREM 2. *Under some parameters $\{\Delta_k\}$ in the PCPS and PCMS models are set to zero, the following equality holds:*

$$G^2(PCPS) = G^2(CPS) + G^2(PCMS).$$

PROOF. We can obtain $\hat{e}_{ij}$ by setting the partial derivatives of Lagrangian equal to zero. Although details are omitted, $\hat{e}_{ij}$ in the PCPS, CPS, and PCMS models are provided as (1), (2), and (3), respectively.

$$\hat{e}_{ij} = \begin{cases} (f_{ij}+f_{ji})\dfrac{\exp\left[\sum_{k=1}^{R-1}(j-1)^{k-1}\hat{\Delta}_k\right]}{1+\exp\left[\sum_{k=1}^{R-1}(j-1)^{k-1}\hat{\Delta}_k\right]} & (i<j), \\ (f_{ij}+f_{ji})\dfrac{1}{1+\exp\left[\sum_{k=1}^{R-1}(j-1)^{k-1}\hat{\Delta}_k\right]} & (i>j), \\ f_{ij} & (i=j), \end{cases} \tag{1}$$

$$\hat{e}_{ij} = \begin{cases} (f_{ij}+f_{ji})\dfrac{f_j^+}{f_j^+ + f_j^-} & (i<j), \\ (f_{ij}+f_{ji})\dfrac{f_i^-}{f_i^+ + f_i^-} & (i>j), \\ f_{ij} & (i=j), \end{cases} \tag{2}$$

$$\hat{e}_{ij} = \begin{cases} f_{ij}\dfrac{f_j^+ + f_j^-}{f_j^+}\dfrac{\exp\left[\sum_{k=1}^{R-1}(j-1)^{k-1}\hat{\Delta}_k\right]}{1+\exp\left[\sum_{k=1}^{R-1}(j-1)^{k-1}\hat{\Delta}_k\right]} & (i<j), \\ f_{ij}\dfrac{f_i^+ + f_i^-}{f_i^-}\dfrac{1}{1+\exp\left[\sum_{k=1}^{R-1}(j-1)^{k-1}\hat{\Delta}_k\right]} & (i>j), \\ f_{ij} & (i=j), \end{cases} \tag{3}$$

where

$$f_i^+ = \sum_{l=1}^{i-1} f_{li} \quad \text{and} \quad f_i^- = \sum_{l=1}^{i-1} f_{il}.$$

In the both PCPS and PCMS models, the $\hat{\Delta}_s$, corresponding Δ_s which is not set to zero, are the solution of following equation

$$\sum_{c=1}^{R-1} \frac{c^{s-1}\left[f_c^+ - f_c^- \exp\left(\sum_{k=1}^{R-1} c^{k-1}\Delta_k\right)\right]}{1+\exp\left(\sum_{k=1}^{R-1} c^{k-1}\Delta_k\right)} = 0.$$

The $f_{ij}/\hat{e}_{ij}$ in the PCPS model is equal to the product of that in the CPS and PCMS models. Therefore, the value of test statistic G^2 of the PCPS model is equal to the sum of that for the CPS and PCMS models. The proof is completed. □

4.2. Model selection

We consider comparing the goodness-of-fit between two nested models for the data. Assume that the model M_2 constantly holds when the model M_1 holds; that is the models M_1 and M_2 is nested. For example, the model M_1 is the PCPS model and M_2 is the CPS model.

Let ν_1 and ν_2 denote the numbers of degrees of freedom for the models M_1 and M_2, respectively. Note that $\nu_1 > \nu_2$ and $G^2(M_1) > G^2(M_2)$. For testing that model M_1 holds assuming that model M_2 holds true, we can use the likelihood ratio statistic $G^2(M_1|M_2) = G^2(M_1) - G^2(M_2)$. When the model M_1 holds, $G^2(M_1|M_2)$ has an asymptotic chi-squared distribution with $\nu_1 - \nu_2$ degrees of freedom.

We consider the PCPS models that are set $\Delta_2 = \Delta_3 = \cdots = \Delta_{R-1} = 0$ and $\Delta_1 = \Delta_3 = \cdots = \Delta_{R-1} = 0$; that is the CS and LCPS models. The relationship between the CS and LCPS models is non-nested. When we compare the goodness-of-fit of all PCPS models for the data, it is necessary to use an index corresponding to compare the non-nested models.

A quick method for choosing the best-fitting model among applied models which include non-nested models is to use the Akaike information criterion (AIC), which is defined as

$$\mathrm{AIC} = -2(\text{maximum log likelihood}) + 2(\text{number of parameters}),$$

for each model (see, e.g., Akaike, 1974). This criterion provides the best-fitting model as the one with minimum AIC. Because only the difference between AICs is required when two models are compared, it is possible to ignore a common constant of AIC, and we may use a modified AIC defined as

$$\mathrm{AIC}^+ = G^2 - 2(\text{number of degrees of freedom}).$$

Thus, the model with the minimum AIC^+ (i.e., the minimum AIC) is the best-fitting model among the applied models. We obtain the following corollary from Theorem 2.

COROLLARY 3. *Under some parameters $\{\Delta_k\}$ in the PCPS and PCMS models are set to zero, the value of AIC^+ for the PCPS model is equal to the sum of that for the CPS and PCMS models.*

It must be noted that Corollary 3 does not hold for the original AIC.

5. APPLICATION TO REAL-WORLD DATA

Consider the data set in Table 1. This data set presents the cross-classification of occupational status categories for father and son dyads in Japan examined in 2015. From Table 2, it is likely that the values of f_{ij}/f_{ji} for all $i < j$ will depend on only the value of column category j (rather than the difference in j and i). Therefore, we forecast that the CPS model fits well for the data set in Table 1. In addition, we are interested in applying the PCPS model which is more parsimonious model than the CPS model to this data set.

Table 3 gives the values of G^2 and AIC^+ for each model. As shown in Table 3, the CPS and PCPS with $\Delta_3 = 0$ models fit well but the other models fit poorly. We compare the goodness-of-fit of the CPS and PCPS with $\Delta_3 = 0$ models using the likelihood ratio statistics shown in Section 4.2. The CPS model is preferable to the PCPS with $\Delta_3 = 0$ model, this is because $G^2(PCPS|CPS) = 5.20$. Moreover, the CPS model is the best-fitting model among the models applied to the data set of Table 1, since the CPS model has the minimum AIC^+.

TABLE 3
The values of the likelihood ratio chi-square statistic G^2 and the modified Akaike's information criterion (AIC^+), for the PCPS model applied to the data set of Table 1.

Parameters $\{\Delta_k\}$ set to zero		Degree of freedom	G^2	AIC^+
$\Delta_1 = \Delta_2 = \Delta_3 = 0$	(S model)	6	129.23*	117.23
$\Delta_2 = \Delta_3 = 0$	(CS model)	5	66.09*	56.09
$\Delta_1 = \Delta_3 = 0$	(LCPS model)	5	37.60*	27.60
$\Delta_1 = \Delta_2 = 0$		5	24.84*	14.84
$\Delta_3 = 0$		4	6.75	−1.25
$\Delta_2 = 0$		4	12.63*	4.63
$\Delta_1 = 0$		4	17.08*	9.08
None	(CPS model)	3	1.55	−4.45

$*$ means significant at the 0.05 level.

Table 4 shows the MLEs of the expected frequencies under the CPS and PCPS with $\Delta_3 = 0$ models.

Under the CPS model, the MLEs of $\exp(\delta_1)$, $\exp(\delta_2)$, and $\exp(\delta_3)$ are $\exp(\hat{\delta}_1) = 5.167$, $\exp(\hat{\delta}_2) = 0.650$, and $\exp(\hat{\delta}_3) = 0.264$. In the CPS model, (i) the odds that an observation will fall in row category i (< 2) and column category 2 instead of row category 2 and column category i are 5.167, (ii) the odds that an observation will fall in row category i (< 3) and column category 3 instead of row category 3 and column category i are 0.650, and (iii) the odds that an observation will fall in row category i (< 4) and column category 4 instead of row category 4 and column category i are 0.264.

Next, using Theorems 1 and 2, for example, we consider the reason the PCPS with $\Delta_1 = 0$ model does not hold. From Theorem 2, the value of G^2 for the PCMS with $\Delta_1 = 0$ model is 15.53 $(= 17.08 - 1.55)$. We can consider that the PCPS with $\Delta_1 = 0$ does not hold because the PCMS with $\Delta_1 = 0$ model does not hold rather than the CPS model.

TABLE 4
The maximum likelihood estimates of expected frequencies under the columns-parameter symmetry (CPS) and polynomial columns-parameter symmetry (PCPS) with $\Delta_3 = 0$ models applied to the data set in Table 1 are shown in parentheses in the second and third lines, respectively.

	Son status				
Father status	(1)	(2)	(3)	(4)	Total
Highest (1)	33	31	10	7	81
	(33)	(31)	(7.48)	(6.47)	
	(33)	(26.36)	(8.25)	(5.95)	
(2)	6	151	81	25	263
	(6)	(151)	(83.52)	(25.47)	
	(10.64)	(151)	(92.02)	(23.40)	
(3)	9	131	207	25	372
	(11.52)	(128.48)	(207)	(25.05)	
	(10.75)	(119.98)	(207)	(23.02)	
Lowest (4)	24	97	95	62	278
	(24.53)	(96.53)	(94.95)	(62)	
	(25.05)	(98.60)	(96.98)	(62)	
Total	72	410	393	119	994

6. Concluding remarks

This study proposed the PCPS model. In the PCPS model, the odds for all $i < j$ that an observation will fall in row category i and column category j instead of row category j and column category i depend on only the value of column category j. This study showed that (i) it is necessary to hold the PCMS model, in addition to the CPS model, to satisfy the PCPS model (see, Theorem 1), and (ii) the value of likelihood ratio chi-square statistic for testing the PCPS model is equal to the sum of that for testing the DPS and PCPS models (see, Theorem 2).

We observe that the CPS model is saturated on the $(1,2)$th and $(2,1)$th cells as well as the main diagonal of the $R \times R$ table, however, the PCPS with $\Delta_3 = 0$ model is saturated only on the main diagonal (see the MLEs of the expected frequencies for the CPS and PCPS with $\Delta_3 = 0$ models in Table 4). For the data set in Table 1, the CPS model fits better than that of the PCPS with $\Delta_3 = 0$ model. However, the PCPS with $\Delta_3 = 0$ model may be preferred over the CPS model when the use of all observations on off-diagonal cells is required.

We suppose that model M_1 holds if and only if both models M_2 and M_3 hold. Thus, the following necessary and sufficient condition holds:

$$U(M_1) \Leftrightarrow U(M_2) \wedge U(M_3),$$

where the number of degrees of freedom for model M_1 is equal to the sum of that for models M_2 and M_3. Darroch and Silvey (1963) described that (i) when the following asymptotic equivalence holds:

$$G^2(M_1) \simeq G^2(M_2) + G^2(M_3), \tag{4}$$

if both models M_2 and M_3 are accepted (at the α significance level) with high probability, then model M_1 would be accepted; however, (ii) when (4) does not hold, it is quite possible for an incompatible situation to arise where both models M_2 and M_3 are accepted with high probability but model M_1 is rejected with high probability (in fact, Darroch and Silvey (1963) and Tahata *et al.* (2011) showed such an interesting example). It must be noted that Theorem 1 would not arise such an incompatible situation.

Acknowledgements

The author would like to thank the anonymous reviewer and the editor for their comments and suggestions to improve this paper.

References

A. Agresti (1983). *A simple diagonals-parameter symmetry and quasi-symmetry model*. Statistics & probability letters, 1, no. 6, pp. 313–316.

H. Akaike (1974). *A new look at the statistical model identification*. IEEE transactions on automatic control, 19, no. 6, pp. 716–723.

S. Ando, K. Aoba (2018). *Decompositions of symmetry using odds-symmetry for square contingency tables*. Journal of Advanced Statistics, 3, no. 3, pp. 23–30.

A. H. Bowker (1948). *A test for symmetry in contingency tables*. Journal of the american statistical association, 43, no. 244, pp. 572–574.

J. Darroch, S. Silvey (1963). *On testing more than one hypothesis*. The Annals of Mathematical Statistics, 34, no. 2, pp. 555–567.

L. A. Goodman (1979). *Multiplicative models for square contingency tables with ordered categories*. Biometrika, 66, no. 3, pp. 413–418.

K. Hashimoto (2018). *Shin Nippon no Kaikyu Shakai (New Japanese Class Society)*. Koudansha Company, Tokyo (in Japanese).

P. McCullagh (1978). *A class of parametric models for the analysis of square contingency tables with ordered categories*. Biometrika, 65, no. 2, pp. 413–418.

K. Tahata, S. Ando, S. Tomizawa (2011). *Ridit score type asymmetry model and decomposition of symmetry for square contingency tables*. Model Assisted Statistics and Applications, 6, no. 4, pp. 279–286.

K. TAHATA, S. TOMIZAWA (2014). *Symmetry and asymmetry models and decompositions of models for contingency tables*. SUT journal of Mathematics, 50, no. 2, pp. 131–165.

S. TOMIZAWA (1984). *Three kinds of decompositions for the conditional symmetry model in a square contingency table*. Journal of the Japan Statistical Society, 14, no. 1, pp. 35–42.

S. TOMIZAWA (1985). *Decompositions for odds-symmetry models in a square contingency table with ordered categories*. Journal of the Japan Statistical Society, 15, no. 2, pp. 151–159.

S. TOMIZAWA (1987). *Diagonal weighted marginal homogeneity model and decompositions for linear diagonals-parameter symmetry model*. Communications in Statistics-Theory and Methods, 16, no. 2, pp. 477–488.

S. TOMIZAWA (1990a). *A decomposition of linear diagonals-parameter symmetry model for square contingency tables with ordered categories*. Biometrical journal, 32, no. 6, pp. 761–766.

S. TOMIZAWA (1990b). *Polynomial diagonals-parameter symmetry model for square contingency tables with ordered categories*. Statistica, 50, no. 2, pp. 171–178.

S. TOMIZAWA (1991). *Decompositions of polynomial diagonals-parameter symmetry models for square contingency tables with ordered categories*. Journal of the Japan Statistical Society, 21, no. 2, pp. 173–180.

S. TOMIZAWA, N. MIYAMOTO, M. IWAMOTO (2006). *Linear column-parameter symmetry model for square contingency tables: application to decayed teeth data*. Biometrical Letters, 43, no. 2, pp. 91–98.

SUMMARY

This study proposes a polynomial columns-parameter symmetry model for square contingency tables with the same row and column ordinal classifications. In the proposed model, the odds for all $i < j$ that an observation will fall in row category i and column category j instead of row category j and column category i depend on only the value of column category j. The proposed model is original because many asymmetry models in square contingency tables depend on the both values of row and column category. The proposed model constantly holds when the columns-parameter symmetry model holds; but the converse does not necessarily hold. This study shows that it is necessary to satisfy the polynomial columns-marginal symmetry model, in addition to the columns-parameter symmetry model, to satisfy the proposed model. This decomposition theorem is useful for explaining why the proposed model does not hold. Moreover, this study shows the value of likelihood ratio chi-square statistic for testing the proposed model is equal to the sum of that for testing the decomposed two models.

Keywords: Asymmetry; Marginal symmetry; Odds; Ordered category; Test statistic.