

Marginal Cumulative Logistic Model of General Order for Multi-way Contingency Tables

Takuya Yoshimoto **Kouji Tahata** **Sadao Tomizawa**
Tokyo University of Science, Tokyo University of Science Tokyo University of Science
Chugai Pharmaceutical Co., Ltd.

Abstract

For multi-way contingency table, [Bhapkar and Darroch \(1990\)](#) considered the marginal symmetry model for order h . The present paper proposes a marginal cumulative logistic model for order h . When $h = 1$, this model reduces to the marginal logistic model ([Agresti 2013](#)). It also gives a theorem that the marginal symmetry model for order h holds if and only if (i) the marginal cumulative logistic model for order h , (ii) the marginal moment equality model for order h , and (iii) the marginal symmetry model for order $h - 1$ hold. A special case of this theorem with $h = 1$ is identical to the result of [Tahata, Katakura, and Tomizawa \(2007\)](#).

Keywords: decomposition, inhomogeneity, marginal symmetry, moment.

1. Introduction

Consider an r^T contingency table ($T \geq 2$). Let X_k ($k = 1, \dots, T$) denote the k th variable. Denote the h th-order ($1 \leq h < T$) marginal probability $\Pr(X_{s_1} = i_1, \dots, X_{s_h} = i_h)$ by $p_{\mathbf{i}}^{\mathbf{s}_h}$, where $\mathbf{s}_h = (s_1, \dots, s_h)$ and $\mathbf{i} = (i_1, \dots, i_h)$ with $1 \leq s_1 < \dots < s_h \leq T$ and $i_k = 1, \dots, r$ ($k = 1, \dots, h$). For a fixed h ($1 \leq h < T$), the h th-order marginal symmetry (M_h^T) model is defined by

$$p_{\mathbf{i}}^{\mathbf{s}_h} = p_{\mathbf{j}}^{\mathbf{s}_h} = p_{\mathbf{i}}^{\mathbf{t}_h},$$

for any permutation $\mathbf{j} = (j_1, \dots, j_h)$ of $\mathbf{i} = (i_1, \dots, i_h)$ and for any $\mathbf{s}_h = (s_1, \dots, s_h)$ and $\mathbf{t}_h = (t_1, \dots, t_h)$ with $1 \leq t_1 < \dots < t_h \leq T$ and $i_k = 1, \dots, r$ ($k = 1, \dots, h$) ([Bhapkar and Darroch 1990](#); [Tomizawa and Tahata 2007](#)). This model indicates the structure of symmetry and homogeneity of h th-order marginal distribution.

For the case of $h = 1$, the M_h^T model is expressed as

$$p_i^{(1)} = \dots = p_i^{(T)} \quad (i = 1, \dots, r),$$

where $p_i^{(k)} = \Pr(X_k = i)$. For instance, see [Stuart \(1955\)](#), [Bishop, Fienberg, and Holland \(1975, p.303\)](#) and [Agresti \(2013, p.439\)](#). This model indicates the homogeneity structure of first-order marginal distribution.

For the multi-way table with *ordinal categories*, several studies considered the marginal cumulative probability in order to discuss the inhomogeneity of first-order marginal distribution. Let $F_i^{(k)}$ denote the first-order marginal cumulative probability and let $L_i^{(k)}$ denote the first-order marginal cumulative logit of X_k for $i = 1, \dots, r-1$, $k = 1, \dots, T$; namely, $F_i^{(k)} = \Pr(X_k \leq i)$ and $L_i^{(k)} = \text{logit}(F_i^{(k)}) = \log(F_i^{(k)} / (1 - F_i^{(k)}))$. The marginal cumulative logistic (ML^T) model is defined by

$$L_i^{(k)} = L_i^{(1)} - \Delta_k \quad (i = 1, \dots, r-1; k = 1, \dots, T),$$

where $\Delta_1 = 0$ (Agresti 2013, p.442). A special case of this model obtained by putting $\{\Delta_k = 0\}$ is the M₁^T model. For instance, when $T = 2$, see McCullagh (1977). Consider the marginal mean equality (ME^T) model defined by

$$E(X_1) = \dots = E(X_T).$$

Agresti (2013, p.440) discussed the decomposition of model. That is, generally suppose that model H₃ implies models H₁ and H₂, model H₃ holds if and only if both models H₁ and H₂ hold. This enables us to see that assuming that model H₁ holds true, the hypothesis that model H₃ holds is equivalent to the hypothesis that the model H₂ holds, and the decomposition of model should be useful to observe the reason for its poor fit when model H₃ does not fit the data well. Tahata *et al.* (2007) noted that for an r^T table, the M₁^T model holds if and only if both the ML^T and ME^T models hold. For $T = 2$, see Miyamoto, Niibe, and Tomizawa (2005). For order h with $1 \leq h < T$, denote the h th-order marginal cumulative probability $\Pr(X_{s_1} \leq i_1, \dots, X_{s_h} \leq i_h)$ by $F_i^{s_h}$, where $s_h = (s_1, \dots, s_h)$ and $\mathbf{i} = (i_1, \dots, i_h)$ with $1 \leq s_1 < \dots < s_h \leq T$ and $i_k = 1, \dots, r$ ($k = 1, \dots, h$). Note that when some i_k equal to r , $F_i^{s_h}$ reduces to the marginal cumulative probability for lower order. For example, when $i_h = r$, $\Pr(X_{s_1} \leq i_1, \dots, X_{s_h} \leq r) = F_i^{s_{h-1}}$. Then, the M_h^T model may be expressed as

$$F_i^{s_h} = F_j^{s_h} = F_i^{t_h},$$

for any permutation $\mathbf{j} = (j_1, \dots, j_h)$ of $\mathbf{i} = (i_1, \dots, i_h)$, where $i_k = 1, \dots, r$ ($k = 1, \dots, h$) and for any $s_h = (s_1, \dots, s_h)$ and $t_h = (t_1, \dots, t_h)$. Since $F_i^{s_h}$ reduces to the marginal cumulative probability for lower order in case where some i_k equal to r , the M_h^T model may also be expressed as

$$F_i^{s_l} = F_j^{s_l} = F_i^{t_l} \quad (l = 1, \dots, h),$$

for any permutation $\mathbf{j} = (j_1, \dots, j_l)$ of $\mathbf{i} = (i_1, \dots, i_l)$, where $i_k = 1, \dots, \overline{r-1}$ ($k = 1, \dots, l$) and for any $s_l = (s_1, \dots, s_l)$ and $t_l = (t_1, \dots, t_l)$. In order to emphasize, hereafter we refer to the $r-1$ as $\overline{r-1}$ in this paper. Note that if the M_h^T model holds, then the M_{h-1}^T model holds, but the converse does not always hold. Hence we are interested in proposing the model we need in order that the M_h^T model holds when the M_{h-1}^T model holds. The ML^T model focuses on the fixed ($h = 1$) order marginal distributions, and describes the inhomogeneity structure based on the logits of $\{F_i^{(k)}\}$. Focusing on the h th-order marginal distribution, we are interested in the symmetry and inhomogeneity structure based on the logits of $\{F_i^{s_h}\}$. Furthermore, we are also interested in the decomposition of the M_h^T model, which is the generalization of the result given by Tahata *et al.* (2007). The decomposition for the M_h^T model should be useful to explore the reason for the poor fit of the M_h^T model when the M_h^T model does not hold for analyzing the data.

In this paper, we propose a marginal cumulative logistic model of general order, and give a decomposition of the M_h^T model by using the proposed model. Section 2 proposes the h th-order marginal cumulative logistic model. Section 3 gives the decompositions of the M_h^T model. Section 4 presents the goodness-of-fit test. Section 5 shows some examples. Finally, Section 6 provides concluding remarks.

2. Models

For a fixed h ($1 \leq h < T$), consider a model defined by

$$L_{\mathbf{i}}^{\mathbf{s}_h} = L_{\mathbf{i}}^{\mathbf{l}_h} - \Delta_{\mathbf{s}_h} \quad \text{and} \quad L_{\mathbf{i}}^{\mathbf{s}_h} = L_{\mathbf{j}}^{\mathbf{s}_h},$$

for any permutation $\mathbf{j} = (j_1, \dots, j_h)$ of $\mathbf{i} = (i_1, \dots, i_h)$ and $\mathbf{l}_h = (1, \dots, h)$, where $i_k = 1, \dots, r-1$ ($k = 1, \dots, h$) and for any $\mathbf{s}_h = (s_1, \dots, s_h)$ with $1 \leq s_1 < \dots < s_h \leq T$,

$$L_{\mathbf{i}}^{\mathbf{s}_h} = \text{logit}(F_{\mathbf{i}}^{\mathbf{s}_h}) = \log\left(\frac{F_{\mathbf{i}}^{\mathbf{s}_h}}{1 - F_{\mathbf{i}}^{\mathbf{s}_h}}\right),$$

where $\Delta_{\mathbf{l}_h} = 0$. We shall refer to this model as the h th-order marginal cumulative logistic (ML $_h^T$) model. This model indicates that the odds of $\{F_{\mathbf{i}}^{\mathbf{l}_h}\}$ is $\exp(\Delta_{\mathbf{s}_h})$ times higher than the odds of $\{F_{\mathbf{i}}^{\mathbf{s}_h}\}$, and the symmetry structure of $\{F_{\mathbf{i}}^{\mathbf{s}_h}\}$ holds. For a fixed h ($1 \leq h < T$), by putting $L_{\mathbf{i}}^{\mathbf{l}_h} = \theta_{\mathbf{i}}$; that is

$$F_{\mathbf{i}}^{\mathbf{l}_h} = \frac{\exp(\theta_{\mathbf{i}})}{1 + \exp(\theta_{\mathbf{i}})},$$

the ML $_h^T$ model can be expressed as a logistic function, for any $\mathbf{s}_h = (s_1, \dots, s_h)$,

$$F_{\mathbf{i}}^{\mathbf{s}_h} = \frac{\exp(\theta_{\mathbf{i}} - \Delta_{\mathbf{s}_h})}{1 + \exp(\theta_{\mathbf{i}} - \Delta_{\mathbf{s}_h})} \quad \text{and} \quad F_{\mathbf{i}}^{\mathbf{s}_h} = F_{\mathbf{j}}^{\mathbf{s}_h},$$

for any permutation $\mathbf{j} = (j_1, \dots, j_h)$ of $\mathbf{i} = (i_1, \dots, i_h)$, where $i_k = 1, \dots, r-1$ ($k = 1, \dots, h$) and $\Delta_{\mathbf{l}_h} = 0$. Therefore this model states that $\{F_{\mathbf{i}}^{\mathbf{l}_h}\}$ is a location shift of $\{F_{\mathbf{i}}^{\mathbf{s}_h}\}$ on a logistic scale. When $h = 1$, the ML $_h^T$ model is identical to the ML T model.

For a fixed h ($1 \leq h < T$), consider a model defined by

$$E(X_{s_1} \cdots X_{s_h}) = E(X_1 \cdots X_h),$$

for $1 \leq s_1 < \dots < s_h \leq T$. We shall refer to this model as the h th-order marginal moment equality (ME $_h^T$) model. When $h = 1$, the ME $_h^T$ model is identical to the ME T model.

3. Decompositions of the marginal symmetry model

Tahata *et al.* (2007) showed the decomposition of the M $_1^T$ model. We shall consider the decomposition of the M $_h^T$ model for an r^T table. Let $X_k^* = r + 1 - X_k$ for $k = 1, \dots, T$. First, we obtain the following lemma.

Lemma 1. For order h ($1 \leq h < T$), when the M $_{h-1}^T$ model holds, the ME $_h^T$ model can be expressed as

$$E(X_{s_1}^* \cdots X_{s_h}^*) = E(X_1^* \cdots X_h^*),$$

for any $1 \leq s_1 < \dots < s_h \leq T$.

Proof.

$$\begin{aligned} E(X_{s_1}^* \cdots X_{s_h}^*) &= (r+1)^h + (-1)^1 (r+1)^{h-1} \sum_{l=1}^h E(X_{s_l}) \\ &\quad + (-1)^2 (r+1)^{h-2} \sum_{l < m} \sum E(X_{s_l} X_{s_m}) + \cdots + (-1)^h E(X_{s_1} \cdots X_{s_h}), \end{aligned}$$

for any $1 \leq s_1 < \dots < s_h \leq T$. Since the M $_{h-1}^T$ model holds, all the ME $_k^T$ models hold for $k = 1, \dots, h-1$. Thus when the M $_{h-1}^T$ model holds, the ME $_h^T$ model is identical to the equation,

$$E(X_{s_1}^* \cdots X_{s_h}^*) = E(X_1^* \cdots X_h^*),$$

for any $1 \leq s_1 < \dots < s_h \leq T$. □

From Lemma 1, we obtain the following theorem.

Theorem 1. For a fixed h ($1 \leq h < T$), the M_h^T model holds if and only if all the ML_h^T , ME_h^T , and M_{h-1}^T models hold, where the M_0^T model indicates the saturated model.

Proof. If the M_h^T model holds, then the ML_h^T , ME_h^T , and M_{h-1}^T models hold. Assuming that for a fixed h , the ML_h^T , ME_h^T , and M_{h-1}^T models hold, then we shall show that the M_h^T model holds. It is easily seen that

$$E(X_1^* \cdots X_h^*) = \sum_{i_1=1}^r \cdots \sum_{i_h=1}^r F_{\mathbf{i}}^{\mathbf{l}_h},$$

where $\mathbf{i} = (i_1, \dots, i_h)$ and $\mathbf{l}_h = (1, \dots, h)$. From Lemma 1, we have

$$\sum_{i_1=1}^r \cdots \sum_{i_h=1}^r (F_{\mathbf{i}}^{\mathbf{s}_h} - F_{\mathbf{i}}^{\mathbf{l}_h}) = 0,$$

for any $\mathbf{s}_h = (s_1, \dots, s_h)$. Since the M_{h-1}^T model holds, the aforementioned equation is also expressed as

$$\sum_{i_1=1}^{r-1} \cdots \sum_{i_h=1}^{r-1} (F_{\mathbf{i}}^{\mathbf{s}_h} - F_{\mathbf{i}}^{\mathbf{l}_h}) = 0.$$

Moreover, since the ML_h^T model holds, we obtain

$$\sum_{i_1=1}^{r-1} \cdots \sum_{i_h=1}^{r-1} \left(\frac{\exp(\theta_{\mathbf{i}} - \Delta_{\mathbf{s}_h})}{1 + \exp(\theta_{\mathbf{i}} - \Delta_{\mathbf{s}_h})} - \frac{\exp(\theta_{\mathbf{i}})}{1 + \exp(\theta_{\mathbf{i}})} \right) = 0,$$

for any $\mathbf{s}_h = (s_1, \dots, s_h)$, where $L_{\mathbf{i}}^{\mathbf{l}_h} = \theta_{\mathbf{i}}$. Thus

$$(\exp(-\Delta_{\mathbf{s}_h}) - 1) \sum_{i_1=1}^{r-1} \cdots \sum_{i_h=1}^{r-1} \frac{\exp(\theta_{\mathbf{i}})}{(1 + \exp(\theta_{\mathbf{i}} - \Delta_{\mathbf{s}_h}))(1 + \exp(\theta_{\mathbf{i}}))} = 0.$$

Since

$$\sum_{i_1=1}^{r-1} \cdots \sum_{i_h=1}^{r-1} \frac{\exp(\theta_{\mathbf{i}})}{(1 + \exp(\theta_{\mathbf{i}} - \Delta_{\mathbf{s}_h}))(1 + \exp(\theta_{\mathbf{i}}))} > 0,$$

we obtain $\Delta_{\mathbf{s}_h} = 0$ for any $\mathbf{s}_h = (s_1, \dots, s_h)$. Namely, we see

$$F_{\mathbf{i}}^{\mathbf{s}_h} = F_{\mathbf{j}}^{\mathbf{s}_h} = F_{\mathbf{i}}^{\mathbf{t}_h},$$

for any permutation $\mathbf{j} = (j_1, \dots, j_h)$ of $\mathbf{i} = (i_1, \dots, i_h)$, where $i_k = 1, \dots, \overline{r-1}$ ($k = 1, \dots, h$) and for any $\mathbf{s}_h = (s_1, \dots, s_h)$ and $\mathbf{t}_h = (t_1, \dots, t_h)$. Since the M_{h-1}^T model holds, we obtain

$$F_{\mathbf{i}}^{\mathbf{s}_h} = F_{\mathbf{j}}^{\mathbf{s}_h} = F_{\mathbf{i}}^{\mathbf{t}_h},$$

for any permutation $\mathbf{j} = (j_1, \dots, j_h)$ of $\mathbf{i} = (i_1, \dots, i_h)$, where $i_k = 1, \dots, r$ ($k = 1, \dots, h$) and for any $\mathbf{s}_h = (s_1, \dots, s_h)$ and $\mathbf{t}_h = (t_1, \dots, t_h)$. Namely, the M_h^T model holds. \square

We note that Theorem 1 is the generalization of the result given by Tahata *et al.* (2007). Also, we obtain the following corollary from Theorem 1.

Corollary 1. The M_{T-1}^T model holds if and only if all the ML_h^T models for $1 \leq h < T$ and all the ME_h^T models for $1 \leq h < T$ hold.

4. Goodness-of-fit test

Let $n_{i_1 \dots i_T}$ denote the observed frequency in the (i_1, \dots, i_T) th cell of the r^T table. Assume that a multinomial distribution is applied to the r^T table. The maximum likelihood estimates (MLEs) of the expected frequencies under each model can be obtained by the Newton-Raphson method in the log-likelihood equations. Each model can be tested for the goodness-of-fit using, for example, the likelihood ratio chi-squared statistic (denoted by G^2) with the corresponding degrees of freedom (df). The test statistic G^2 for model H is given by

$$G^2(\text{H}) = 2 \sum_{i_1=1}^r \cdots \sum_{i_T=1}^r n_{i_1 \dots i_T} \log \left(\frac{n_{i_1 \dots i_T}}{\hat{m}_{i_1 \dots i_T}} \right),$$

where $\hat{m}_{i_1 \dots i_T}$ is the MLE of expected frequency $m_{i_1 \dots i_T}$ under model H. Table 1 lists df for each model. We note that the number of df for the M_h^T model is equal to the sum of those for the decomposed models.

Table 1: Numbers of degrees of freedom for models applied to the r^T table.

Models	Degrees of freedom
M_h^T	$\sum_{u=0}^h \binom{T}{u} (r-1)^u - \sum_{u=0}^h \binom{r-2+u}{u}$
ML_h^T	$\left[(r-1)^h - \binom{r-2+h}{h} \right] + \left[\binom{T}{h} - 1 \right] \left[(r-1)^h - 1 \right]$
ME_h^T	$\binom{T}{h} - 1$

[Akaike \(1974\)](#) information criterion (AIC) is used to choose the preferable model among different models which include non-nested models. For details see [Konishi and Kitagawa \(2008\)](#). Since only the difference between AIC's is required when two models are compared, it is possible to ignore a common constant of AIC. We may use a modified AIC defined by

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

Thus, for the data, the model with the minimum AIC^+ is the preferable model. This criterion will be used in the next section.

5. Examples

Consider the data in Tables 2 and 3, taken from the 2016 General Social Survey ([Smith, Davern, Freese, and Hout 2018](#)) conducted by the National Opinion Research Center at the University of Chicago. These describe the cross classifications of subject's opinions regarding government spending on Education (X_1), Environment (X_2), and Assistance to the poor (X_3) in 1984 and 2016, respectively. The common response categories are (1) 'too little', (2) 'about right', and (3) 'too much'.

5.1. Analysis of Table 2

Table 4 gives the values of G^2 and AIC^+ for the data in Table 2, and shows that all models fit the data well since these models are accepted at the 0.05 significance level. Since these

Table 2: Opinions about government spending in 1984 from the 2016 General Social Survey (Smith *et al.* 2018). (Parenthesized values are the maximum likelihood estimates of the expected frequencies under the ML_2^3 model.)

Education (X_1)	Environment (X_2)	Assistance to the poor (X_3)		
		(1) too little	(2) about right	(3) too much
(1) too little	(1) too little	152 (152.06)	34 (38.60)	14 (12.28)
	(2) about right	45 (34.94)	20 (21.98)	8 (5.72)
	(3) too much	19 (17.93)	2 (2.00)	2 (2.02)
(2) about right	(1) too little	34 (40.51)	19 (18.54)	4 (3.58)
	(2) about right	18 (16.65)	26 (25.81)	7 (5.26)
	(3) too much	5 (5.42)	3 (2.57)	2 (2.02)
(3) too much	(1) too little	4 (5.14)	4 (5.59)	5 (5.06)
	(2) about right	9 (11.15)	1 (2.06)	6 (6.07)
	(3) too much	2 (2.02)	2 (2.02)	1 (1.01)

Table 3: Opinions about government spending in 2016 from the 2016 General Social Survey (Smith *et al.* 2018). (Parenthesized values are the maximum likelihood estimates of the expected frequencies under the ML_1^3 model.)

Education (X_1)	Environment (X_2)	Assistance to the poor (X_3)		
		(1) too little	(2) about right	(3) too much
(1) too little	(1) too little	612 (612.35)	110 (106.04)	30 (32.37)
	(2) about right	134 (133.13)	55 (52.66)	11 (11.78)
	(3) too much	51 (51.61)	11 (10.72)	11 (12.02)
(2) about right	(1) too little	85 (89.46)	30 (30.36)	6 (6.84)
	(2) about right	46 (48.05)	43 (43.21)	9 (10.17)
	(3) too much	9 (9.59)	11 (11.26)	5 (5.77)
(3) too much	(1) too little	12 (10.88)	8 (7.01)	3 (2.91)
	(2) about right	16 (14.41)	16 (13.94)	8 (7.71)
	(3) too much	13 (11.91)	8 (7.08)	13 (12.76)

models are including non-nested models, we use AIC^+ to choose the preferable model. Since the ML_2^3 model has a minimum AIC^+ value, the ML_2^3 model is the preferable model among the models. Thus it is inferred that there is a symmetry structure but not homogeneity for second-order marginal distribution. With regard to the inhomogeneity structure, the MLEs of parameters $\exp(\Delta_{(1,3)})$ and $\exp(\Delta_{(2,3)})$ under the ML_2^3 model are $\exp(\hat{\Delta}_{(1,3)}) = 0.98$ and $\exp(\hat{\Delta}_{(2,3)}) = 1.05$. For the inhomogeneity structure, the ML_2^3 model provides, for example, the odds that the opinions of education and the environment are both ‘too little’ instead of neither ‘too little’ are estimated to be 0.98 times higher than the odds that the opinions of education and assistance to the poor are both ‘too little’ instead of neither ‘too little’. We can obtain similar results although the details are omitted. This indicates that there is location shift on a logistic scale between the marginal distribution for the opinions of education and the environment and that of education and assistance to the poor. With regard to the case of the environment and assistance to the poor, it can be interpreted in the same way to the case of above.

Table 4: Values of the likelihood ratio chi-squared statistic G^2 and AIC^+ for the models applied to Table 2.

Applied models	Degrees of freedom	G^2	AIC^+
M_2^3	13	21.08	-4.92
ML_2^3	7	8.30	-5.70 ^(min)
ME_2^3	2	0.70	-3.30
M_1^3	4	9.40	1.40
ML_1^3	2	5.49	1.49
ME_1^3	2	3.47	-0.53

(min) means a minimum AIC^+ .

5.2. Analysis of Table 3

Table 5 gives the values of G^2 and AIC^+ for models applied to the data in Table 3. These show that the ML_1^3 model fits the data well although the other models fit the data poorly. We see that the poor fit of the M_2^3 model is due to the lack of the ME_2^3 , ME_1^3 , and ML_2^3 models rather than the ML_1^3 model from Corollary 1. Therefore, it is inferred that the poor fit of the M_2^3 model is caused by the influence of the lack of structure of (i) the equality of the second-order moments of (X_1, X_2) , (X_1, X_3) and (X_2, X_3) , (ii) the equality of means of X_1 , X_2 and X_3 , and (iii) the ML_2^3 model. Under the ML_1^3 , the MLEs of parameters $\exp(\Delta_{(2)})$ and $\exp(\Delta_{(3)})$ are $\exp(\hat{\Delta}_{(2)}) = 1.55$ and $\exp(\hat{\Delta}_{(3)}) = 1.17$. Thus, under the ML_1^3 model, the odds that the opinion is ‘too little’ instead of not ‘too little’ are estimated to be 1.55 times higher in education than in the environment. In a similar manner, we can see that the odds for the opinion of education is also estimated to be 1.55 times higher than that of the environment in either cases. Furthermore, we can interpret that there is location shift on a logistic scale between the marginal distribution for the opinions of education and that of the environment. With regard to the case of the education and assistance to the poor, it can be interpreted in the same way to the case of above.

6. Concluding remarks

In this paper, (i) we have proposed the ML_h^T model, and (ii) given the decomposition of the M_h^T model. The ML_h^T model is the extension of the ML^T model discussed by Agresti (2013, p.442), and the decomposition by using the ML_h^T model is the generalization of the result given by Tahata *et al.* (2007). The decomposition for the M_h^T model should be useful to explore the reason for the poor fit of the M_h^T model when the M_h^T model does not hold for

Table 5: Values of the likelihood ratio chi-squared statistic G^2 and AIC^+ for the models applied to Table 3.

Applied models	Degrees of freedom	G^2	AIC^+
M_2^3	13	57.95*	31.95
ML_2^3	7	25.99*	11.99
ME_2^3	2	28.47*	24.47
M_1^3	4	43.41*	35.41
ML_1^3	2	2.24	-1.76 ^(min)
ME_1^3	2	38.20*	34.20

*significant at the 0.05 level.

(*min*) means a minimum AIC^+ .

analyzing the data. Meanwhile, Theorem 1 also leads to Corollary 1 that decomposes for the M_h^T model into more models. The decomposition for the M_h^T model into more (three or four) models rather than into two models would be useful for exploring the reason for the poor fit in more details when the M_h^T model does not fit well. In practice, Corollary 1 reveals the origin of the poor fit of the M_2^3 model (see Section 5.2.).

Acknowledgements

The authors are greatly thankful to two referees and editor for their helpful and valuable suggestions to further improve this article.

References

- Agresti A (2013). *Categorical Data Analysis, 3rd edition*. New Jersey/Wiley.
- Akaike H (1974). “A New Look at the Statistical Model Identification.” *IEEE Transactions on Automatic Control*, **AC-19**, 716–723. doi:10.1109/TAC.1974.1100705.
- Bhapkar VP, Darroch JN (1990). “Marginal Symmetry and Quasi Symmetry of General Order.” *Journal of Multivariate Analysis*, **34**, 173–184. doi:10.1016/0047-259X(90)90034-F.
- Bishop YMM, Fienberg SE, Holland PW (1975). *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Massachusetts/The MIT Press.
- Konishi S, Kitagawa G (2008). *Information Criteria and Statistical Modeling*. New York/Springer. doi:10.1007/978-0-387-71887-3.
- McCullagh P (1977). “A Logistic Model for Paired Comparisons with Ordered Categorical Data.” *Biometrika*, **64**, 449–453. doi:10.1093/biomet/64.3.449.
- Miyamoto N, Niibe K, Tomizawa S (2005). “Decompositions of Marginal Homogeneity Model Using Cumulative Logistic Models for Square Contingency Tables with Ordered Categories.” *Austrian Journal of Statistics*, **34**, 361–373. doi:10.17713/ajs.v34i4.424.
- Smith TW, Davern M, Freese J, Hout M (2018). “General Social Surveys, 1972-2016 [machine-readable data file] /Principal Investigator, Smith, T. W.; Co-Principal Investigator, Davern, M.; Co-Principal Investigator, Freese, J.; Co-Principal Investigator, Hout, M.; Sponsored by National Science Foundation. –NORC ed.– Chicago: NORC at the University of Chiago

[producer]; Storrs, CT: The Roper Center for Public Opinion Research, University of Connecticut [distributor]. 1 data file (62,466 logical records) + 1 codebook (3,689.). – (National Data Program for the Social Sciences, No. 24).”

Stuart A (1955). “A Test for Homogeneity of the Marginal Distributions in a Two-way Classification.” *Biometrika*, **42**, 412–416. doi:10.1093/biomet/42.3-4.412.

Tahata K, Katakura S, Tomizawa S (2007). “Decompositions of Marginal Homogeneity Model Using Cumulative Logistic Models for Multi-way Contingency Tables.” *Revstat: Statistical Journal*, **5**, 163–176.

Tomizawa S, Tahata K (2007). “The Analysis of Symmetry and Asymmetry: Orthogonality of Decomposition of Symmetry into Quasi-symmetry and Marginal Symmetry for Multi-way Tables.” *Journal de la Société Française de Statistique*, **148**, 3–36.

Affiliation:

Takuya Yoshimoto
Department of Information Science
Faculty of Science and Technology
Tokyo University of Science
Noda City, Chiba, Japan

Clinical Information & Intelligence Dept.
Chugai Pharmaceutical Co., Ltd.
Chuo-ku, Tokyo, Japan
E-mail: yoshimoto.takuya61@chugai-pharm.co.jp

Kouji Tahata
Department of Information Science
Faculty of Science and Technology
Tokyo University of Science
Noda City, Chiba, Japan
E-mail: kouji_tahata@is.noda.tus.ac.jp

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