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

Takuya Yoshimoto  
Tokyo University of Science

Kouji Tahata  
Tokyo University of Science

Sadao Tomizawa  
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 $X_k$ ($k = 1, \ldots, T$) denote the $k$th variable. Denote the $h$th-order ($1 \leq h < T$) marginal probability $\Pr(X_{s_1} = i_1, \ldots, X_{s_h} = i_h)$ by $p_{i_1}^{s_1} \cdots p_{i_h}^{s_h}$, where $s_h = (s_1, \ldots, s_h)$ and $i = (i_1, \ldots, i_h)$ with $1 \leq s_1 < \cdots < s_h \leq T$ and $i_k = 1, \ldots, r$ ($k = 1, \ldots, h$). For a fixed $h$ ($1 \leq h < T$), the $h$th-order marginal symmetry ($M_T^h$) model is defined by

$$p_{i_1}^{s_1} \cdots p_{i_h}^{s_h} = p_{j_1}^{s_1} \cdots p_{j_h}^{s_h} = p_{i_1}^{t_1} \cdots p_{i_h}^{t_h},$$

for any permutation $j = (j_1, \ldots, j_h)$ of $i = (i_1, \ldots, i_h)$ and for any $s_h = (s_1, \ldots, s_h)$ and $t_h = (t_1, \ldots, t_h)$ with $1 \leq t_1 < \cdots < t_h \leq T$ and $i_k = 1, \ldots, r$ ($k = 1, \ldots, h$) (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 $M_T^1$ model is expressed as

$$p_{i_1}^{(1)} = \cdots = p_{i}^{(T)} \quad (i = 1, \ldots, r),$$

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 $i = 1, \ldots, r - 1, k = 1, \ldots, T$; namely, $F_i^{(k)} = \Pr(X_k \leq i)$ and $L_i^{(k)} = \logit \left( F_i^{(k)} \right) = \log \left( F_i^{(k)}/(1 - F_i^{(k)}) \right)$. The marginal cumulative logistic (ML) model is defined by

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

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

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

Agresti (2013, p.440) discussed the decomposition of model. That is, generally suppose that model $H_3$ implies models $H_1$ and $H_2$, model $H_3$ holds if and only if both models $H_1$ and $H_2$ hold. This enables us to see that assuming that model $H_1$ holds true, the hypothesis that model $H_3$ holds is equivalent to the hypothesis that the model $H_2$ holds, and the decomposition of model should be useful to observe the reason for its poor fit when model $H_3$ does not fit the data well. Tahata et al. (2007) noted that for an $r^T$ table, the $M_T^T$ 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 $\Pr(X_{i_1} \leq s_{i_1}, \ldots, X_{i_h} \leq s_{i_h})$ by $F_i^{s_h}$, where $s_h = (s_1, \ldots, s_h)$ and $i = (i_1, \ldots, i_h)$ with $1 \leq s_1 < \cdots < s_h \leq T$ and $i_k = 1, \ldots, r$ ($k = 1, \ldots, h$). 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 $i_h = r$, $\Pr(X_{s_1} \leq i_1, \ldots, X_{s_h} \leq r) = F_i^{s_{h-1}}$. Then, the $M_T^h$ model may be expressed as

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

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

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

for any permutation $j = (j_1, \ldots, j_l)$ of $i = (i_1, \ldots, i_l)$, where $i_k = 1, \ldots, r - 1$ ($k = 1, \ldots, l$) and for any $s_l = (s_1, \ldots, s_l)$ and $t_l = (t_1, \ldots, t_l)$. In order to emphasize, hereafter we refer to the $r - 1$ as $r - 1$ in this paper. Note that if the $M_T^h$ model holds, then the $M_T^{h-1}$ model holds, but the converse does not always hold. Hence we are interested in proposing the model we need in order that the $M_T^h$ model holds when the $M_T^{h-1}$ 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 $M_T^h$ model, which is the generalization of the result given by Tahata et al. (2007). The decomposition for the $M_T^h$ model should be useful to explore the reason for the poor fit of the $M_T^h$ model when the $M_T^h$ 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 $M_T^h$ model by using the proposed model. Section 2 proposes the $h$th-order marginal cumulative logistic model. Section 3 gives the decompositions of the $M_T^h$ 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 \(1 \leq h < T\), consider a model defined by

\[ L^h_i = L^h_i - \Delta^h_i \quad \text{and} \quad L^{s^h}_i = L^{s^h}_j, \]

for any permutation \(j = (j_1, \ldots, j_h)\) of \(i = (i_1, \ldots, i_h)\) and \(l_h = (1, \ldots, h)\), where \(i_k = 1, \ldots, r - 1\) \((k = 1, \ldots, h)\) and for any \(s_h = (s_1, \ldots, s_h)\) with \(1 \leq s_1 < \cdots < s_h \leq T\),

\[ L^{s^h}_i = \logit \left( F^{s^h}_i \right) = \log \left( \frac{F_i^{s^h}}{1 - F_i^{s^h}} \right), \]

where \(\Delta^h_i = 0\). We shall refer to this model as the \(h\)th-order marginal cumulative logistic (ML\(^T\)) model. This model indicates that the odds of \(F^{s^h}_i\) is \(\exp(\Delta^h_i)\) times higher than the odds of \(F^{s^h}_i\), and the symmetry structure of \(F^{s^h}_i\) holds. For a fixed \(1 \leq h < T\), by putting \(L^h_i = \theta_i\); that is

\[ F^{h^h}_i = \frac{\exp(\theta_i)}{1 + \exp(\theta_i)}, \]

the ML\(^T\) model can be expressed as a logistic function, for any \(s_h = (s_1, \ldots, s_h)\),

\[ F^{s^h}_i = \frac{\exp(\theta_i - \Delta^h_i)}{1 + \exp(\theta_i - \Delta^h_i)} \quad \text{and} \quad F^{s^h}_j = F^{s^h}_j, \]

for any permutation \(j = (j_1, \ldots, j_h)\) of \(i = (i_1, \ldots, i_h)\), where \(i_k = 1, \ldots, r - 1\) \((k = 1, \ldots, h)\) and \(\Delta^h_i = 0\). Therefore this model states that \(F^{s^h}_i\) is a location shift of \(F^{s^h}_i\) on a logistic scale. When \(h = 1\), the ML\(^T\) model is identical to the ML\(^T\) model.

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

\[ E(X_{s^1} \cdots X_{s^h}) = E(X_1 \cdots X_h), \]

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

3. Decompositions of the marginal symmetry model

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

**Lemma 1.** For order \(1 \leq h < T\), when the ME\(^T\) model holds, the ME\(^T\) model can be expressed as

\[ E(X_{s^1}^* \cdots X_{s^h}^*) = E(X_1^* \cdots X_h^*), \]

for any \(1 \leq s_1 < \cdots < s_h \leq T\).

**Proof.**

\[
E(X_{s^1}^* \cdots X_{s^h}^*) = (r + 1)^h + (-1)^1(r + 1)^{h-1} \sum_{l=1}^{h} E(X_{s_l}) \\
+ (-1)^2(r + 1)^{h-2} \sum_{l<m} E(X_{s_l}X_{s_m}) + \cdots + (-1)^h E(X_{s_1} \cdots X_{s_h}),
\]

for any \(1 \leq s_1 < \cdots < s_h \leq T\). Since the ME\(^T\) model holds, all the ME\(^T\) models hold for \(k = 1, \ldots, h - 1\). Thus when the ME\(^T\) model holds, the ME\(^T\) model is identical to the equation,

\[ E(X_{s^1}^* \cdots X_{s^h}^*) = E(X_1^* \cdots X_h^*), \]

for any \(1 \leq s_1 < \cdots < s_h \leq T\).  

\[\square\]
Moreover, since the ML model holds, all the Corollary 1.
Theorem 1. For a fixed \( h \) (\( 1 \leq h < T \)), the \( M^T_h \) model holds if and only if all the \( ML^T_h, ME^T_h, \) and \( M^T_{h-1} \) models hold, where the \( M^0_h \) model indicates the saturated model.

Proof. If the \( M^T_h \) model holds, then the \( ML^T_h, ME^T_h, \) and \( M^T_{h-1} \) models hold. Assuming that for a fixed \( h \), the \( ML^T_h, ME^T_h, \) and \( M^T_{h-1} \) models hold, then we shall show that the \( M^T_h \) model holds. It is easily seen that

\[
E(X^*_1 \cdots X^*_h) = \sum_{i_1=1}^{r} \cdots \sum_{i_h=1}^{r} F^t_{i_1} \cdots F^t_{i_h},
\]

where \( i = (i_1, \ldots, i_h) \) and \( l_h = (1, \ldots, h) \). From Lemma 1, we have

\[
\sum_{i_1=1}^{r} \cdots \sum_{i_h=1}^{r} (F^{s_h}_{i_1} - F^t_{i_1}) = 0,
\]

for any \( s_h = (s_1, \ldots, s_h) \). Since the \( M^T_{h-1} \) model holds, the aforementioned equation is also expressed as

\[
\sum_{i_1=1}^{r-1} \cdots \sum_{i_h=1}^{r-1} (F^{s_h}_{i_1} - F^t_{i_1}) = 0.
\]

Moreover, since the \( ML^T_h \) model holds, we obtain

\[
\sum_{i_1=1}^{r-1} \cdots \sum_{i_h=1}^{r-1} \left( \frac{\exp(\theta_i - \Delta_{s_h})}{1 + \exp(\theta_i - \Delta_{s_h})} - \frac{\exp(\theta_i)}{1 + \exp(\theta_i)} \right) = 0,
\]

for any \( s_h = (s_1, \ldots, s_h) \), where \( L^t_{i_1} = \theta_i \). Thus

\[
(\exp(-\Delta_{s_h}) - 1) \sum_{i_1=1}^{r-1} \cdots \sum_{i_h=1}^{r-1} \frac{\exp(\theta_i)}{(1 + \exp(\theta_i - \Delta_{s_h}))(1 + \exp(\theta_i))} = 0.
\]

Since

\[
\sum_{i_1=1}^{r-1} \cdots \sum_{i_h=1}^{r-1} \frac{\exp(\theta_i)}{(1 + \exp(\theta_i - \Delta_{s_h}))(1 + \exp(\theta_i))} > 0,
\]

we obtain \( \Delta_{s_h} = 0 \) for any \( s_h = (s_1, \ldots, s_h) \). Namely, we see

\[
F^{s_h}_{i_1} = F^{s_h}_{j_1} = F^t_{i_1},
\]

for any permutation \( j = (j_1, \ldots, j_h) \) of \( i = (i_1, \ldots, i_h) \), where \( i_k = 1, \ldots, r - 1 \) (\( k = 1, \ldots, h \)) and for any \( s_h = (s_1, \ldots, s_h) \) and \( t_h = (t_1, \ldots, t_h) \). Since the \( M^T_{h-1} \) model holds, we obtain

\[
F^{s_h}_{i_1} = F^{s_h}_{j_1} = F^t_{i_1},
\]

for any permutation \( j = (j_1, \ldots, j_h) \) of \( i = (i_1, \ldots, i_h) \), where \( i_k = 1, \ldots, r \) (\( k = 1, \ldots, h \)) and for any \( s_h = (s_1, \ldots, s_h) \) and \( t_h = (t_1, \ldots, t_h) \). Namely, the \( M^T_h \) model holds.

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 \( M^T_{h-1} \) model holds if and only if all the \( ML^T_{h} \) models for \( 1 \leq h < T \) and all the \( ME^T_{h} \) models for \( 1 \leq h < T \) hold.
4. Goodness-of-fit test

Let \( n_{i_1 \ldots i_T} \) denote the observed frequency in the \((i_1, \ldots, i_T)\)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

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

where \(\hat{m}_{i_1 \ldots i_T}\) is the MLE of expected frequency \(m_{i_1 \ldots i_T}\) under model H. Table 1 lists df for each model. We note that the number of df for the \(M^T_h\) 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.

<table>
<thead>
<tr>
<th>Models</th>
<th>Degrees of freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>(M^T_h)</td>
<td>(\sum_{u=0}^{h} \left( \frac{T}{u} \right) (r-1)^u - \sum_{u=0}^{h} \left( \frac{r-2+u}{u} \right))</td>
</tr>
<tr>
<td>(ML^T_h)</td>
<td>(\left[ (r-1)^h - \left( \frac{r-2+h}{h} \right) \right] + \left[ \left( \frac{T}{h} \right) - 1 \right] \left( (r-1)^h - 1 \right))</td>
</tr>
<tr>
<td>(ME^T_h)</td>
<td>(\left( \frac{T}{h} \right) - 1)</td>
</tr>
</tbody>
</table>

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

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

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$_3^2$ model.)

<table>
<thead>
<tr>
<th>Education ($X_1$)</th>
<th>Environment ($X_2$)</th>
<th>Assistance to the poor ($X_3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) too little</td>
<td>(1) too little</td>
<td>(1) too little</td>
</tr>
<tr>
<td></td>
<td></td>
<td>152</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(152.06)</td>
</tr>
<tr>
<td>(1) too little</td>
<td>(2) about right</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(34.94)</td>
</tr>
<tr>
<td>(1) too little</td>
<td>(3) too much</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(17.93)</td>
</tr>
<tr>
<td>(2) about right</td>
<td>(1) too little</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(40.51)</td>
</tr>
<tr>
<td>(2) about right</td>
<td>(2) about right</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(16.65)</td>
</tr>
<tr>
<td>(2) about right</td>
<td>(3) too much</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5.42)</td>
</tr>
<tr>
<td>(3) too much</td>
<td>(1) too little</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5.14)</td>
</tr>
<tr>
<td>(3) too much</td>
<td>(2) about right</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(11.15)</td>
</tr>
<tr>
<td>(3) too much</td>
<td>(3) too much</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.02)</td>
</tr>
</tbody>
</table>

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$_3^3$ model.)

<table>
<thead>
<tr>
<th>Education ($X_1$)</th>
<th>Environment ($X_2$)</th>
<th>Assistance to the poor ($X_3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) too little</td>
<td>(1) too little</td>
<td>(1) too little</td>
</tr>
<tr>
<td></td>
<td></td>
<td>612</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(612.35)</td>
</tr>
<tr>
<td>(1) too little</td>
<td>(2) about right</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(133.13)</td>
</tr>
<tr>
<td>(1) too little</td>
<td>(3) too much</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(51.61)</td>
</tr>
<tr>
<td>(2) about right</td>
<td>(1) too little</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(89.46)</td>
</tr>
<tr>
<td>(2) about right</td>
<td>(2) about right</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(48.05)</td>
</tr>
<tr>
<td>(2) about right</td>
<td>(3) too much</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(9.59)</td>
</tr>
<tr>
<td>(3) too much</td>
<td>(1) too little</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(10.88)</td>
</tr>
<tr>
<td>(3) too much</td>
<td>(2) about right</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(14.41)</td>
</tr>
<tr>
<td>(3) too much</td>
<td>(3) too much</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(11.91)</td>
</tr>
</tbody>
</table>
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>
<thead>
<tr>
<th>Applied models</th>
<th>Degrees of freedom</th>
<th>$G^2$</th>
<th>AIC$^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>M$_2^3$</td>
<td>13</td>
<td>21.08</td>
<td>−4.92</td>
</tr>
<tr>
<td>ML$_3^3$</td>
<td>7</td>
<td>8.30</td>
<td>−5.70$^{(\text{min})}$</td>
</tr>
<tr>
<td>ME$_2^3$</td>
<td>2</td>
<td>0.70</td>
<td>−3.30</td>
</tr>
<tr>
<td>M$_3^1$</td>
<td>4</td>
<td>9.40</td>
<td>1.40</td>
</tr>
<tr>
<td>ML$_1^3$</td>
<td>2</td>
<td>5.49</td>
<td>1.49</td>
</tr>
<tr>
<td>ME$_1^3$</td>
<td>2</td>
<td>3.47</td>
<td>−0.53</td>
</tr>
</tbody>
</table>

$^{(\text{min})}$ 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$_3^2$ 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$_3^2$ 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 $\exp(\Delta_{2,3})$ and $\exp(\hat{\Delta}_{1,3})$ are $\exp(\hat{\Delta}_{2,3}) = 1.55$ and $\exp(\hat{\Delta}_{1,3}) = 1.17$. 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$_h^T$ 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.

<table>
<thead>
<tr>
<th>Applied models</th>
<th>Degrees of freedom</th>
<th>$G^2$</th>
<th>AIC$^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_2^3$</td>
<td>13</td>
<td>57.95*</td>
<td>31.95</td>
</tr>
<tr>
<td>$ML_2^3$</td>
<td>7</td>
<td>25.99*</td>
<td>11.99</td>
</tr>
<tr>
<td>$ME_2^3$</td>
<td>2</td>
<td>28.47*</td>
<td>24.47</td>
</tr>
<tr>
<td>$M_1^4$</td>
<td>4</td>
<td>43.41*</td>
<td>35.41</td>
</tr>
<tr>
<td>$ML_1^4$</td>
<td>2</td>
<td>2.24</td>
<td>$-1.76^{(min)}$</td>
</tr>
<tr>
<td>$ME_1^4$</td>
<td>2</td>
<td>38.20*</td>
<td>34.20</td>
</tr>
</tbody>
</table>

*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 $M_2^3$ model into more models. The decomposition for the $M_1^4$ 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 $M_2^3$ 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


**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