SUT Journal of Mathematics
Vol. 55, No. 1 (2019), 25–37

Extended marginal homogeneity models based on complementary log-log transform for multi-way contingency tables

Satoru Shinoda, Kouji Tahata, Kiyotaka Iki and Sadao Tomizawa

(Received September 26, 2018; Revised February 7, 2019)

Abstract. For square contingency tables with ordered categories, Saigusa *et al.* (2018) proposed the marginal cumulative complementary log-log model being an extension of the marginal homogeneity model. The present paper considers the marginal cumulative complementary log-log and conditional marginal cumulative complementary log-log models for multi-way tables. It also gives the decompositions of the marginal homogeneity model into the proposed model and a model of the equality of marginal means for multi-way tables. An example is given.

AMS 2010 *Mathematics Subject Classification.* 62H17.

Key words and phrases. Decomposition, marginal homogeneity, mean equality, ordered categories.

§1. Introduction

For an $R \times R$ square contingency table with ordered categories, let p_{ij} denote the probability that an observation will fall in the cell in row i and column j ($i = 1, \ldots, R;\ j = 1, \ldots, R$), and let X_1 and X_2 denote the row and column variables, respectively. The marginal homogeneity (MH) model is defined by

$$\Pr(X_1 = i) = \Pr(X_2 = i) \quad \text{for } i = 1, \ldots, R;$$

that is

$$p_{i\cdot} = p_{\cdot i} \quad \text{for } i = 1, \ldots, R,$$

where $p_{i\cdot} = \sum_{k=1}^{R} p_{ik}$ and $p_{\cdot i} = \sum_{k=1}^{R} p_{ki}$. This model indicates that the row marginal distribution is identical to the column marginal distribution (Stuart, 1955; Bhapkar, 1966; Bishop, Fienberg and Holland, 1975, p.294). Some extensions of the MH model were proposed by e.g., Agresti (1984, p.205), and Miyamoto, Niibe and Tomizawa (2005).

Let $F_i^{(1)}$ and $F_i^{(2)}$ denote the marginal cumulative probabilities of X_1 and X_2, respectively, i.e., $F_i^{(1)} = \Pr(X_1 \le i) = \sum_{k=1}^{i} p_{k\cdot}$ and $F_i^{(2)} = \Pr(X_2 \le i) = \sum_{k=1}^{i} p_{\cdot k}$ for $i = 1, \dots, R-1$. Then the MH model may be expressed as

$$F_i^{(1)} = F_i^{(2)} \quad \text{for } i = 1, \dots, R-1.$$

Let $C_i^{(1)}$ and $C_i^{(2)}$ denote the marginal cumulative complementary log-log transform of X_1 and X_2, respectively; namely

$$C_i^{(1)} = \log\left(-\log\left(1 - F_i^{(1)}\right)\right),$$
$$C_i^{(2)} = \log\left(-\log\left(1 - F_i^{(2)}\right)\right),$$

for $i = 1, \dots, R-1$. Then the MH model may also be expressed as

$$C_i^{(1)} = C_i^{(2)} \quad \text{for } i = 1, \dots, R-1.$$

Saigusa, Maruyama, Tahata and Tomizawa (2018) proposed the marginal cumulative complementary log-log (MCL) model defined by

$$C_i^{(1)} = C_i^{(2)} + \log \Delta \quad \text{for } i = 1, \dots, R-1,$$

where the parameter Δ is unspecified. The MCL model states that one marginal distribution is a location shift of the other marginal distribution on a complementary log-log scale. A special case of the MCL model obtained by putting $\Delta = 1$ is the MH model.

Consider a specified monotonic function $g(k)$ satisfying $g(1) \le \cdots \le g(R)$ or $g(1) \ge \cdots \ge g(R)$, where at least one strict inequality holds. The marginal mean equality (ME) model is defined by

$$\sum_{i=1}^{R} g(i) p_{i\cdot} = \sum_{i=1}^{R} g(i) p_{\cdot i} \quad (\text{i.e., } \mathrm{E}(g(X_1)) = \mathrm{E}(g(X_2))).$$

Saigusa *et al*. (2018) stated that the MH model holds if and only if both the MCL and ME models hold.

Consider a multi-way R^T contingency table ($T \ge 2$). The MH model for R^T table was given by e.g., Bishop *et al*., 1975, p.303; Bhapkar and Darroch, 1990; Agresti, 2002, p.440. Some extensions of the MH model were proposed by e.g., McCullagh (1977), Tahata, Katakura and Tomizawa (2007), and Tahata, Kobayashi and Tomizawa (2008).

The purpose of the present paper is to extend the MCL model into the R^T table, and to give a decomposition of the MH model for the R^T table. The MH model does not depend on the main diagonal cell probabilities, however, the MCL model depends on them. We are also interested in proposing the other MCL model which does not depend on the main diagonal cell probabilities, namely, in the conditional MCL model on condition that an observation will fall in one of off-diagonal cells of the table. Also, we give a new decomposition of the MH model using the conditional MCL model.

§2. Models

2.1. Marginal cumulative complementary log-log model

Consider an R^T table ($T \geq 2$) having ordered categories. Let X_t denote the t-th random variable for $t = 1, \ldots, T$ and let $\Pr(X_1 = i_1, \ldots, X_T = i_T) = p_{i_1 \ldots i_T}$ for $i_t = 1, \ldots, R$. The MH model is defined by

$$\Pr(X_1 = i) = \cdots = \Pr(X_T = i) \quad \text{for } i = 1, \ldots, R;$$

that is

$$p_i^{(1)} = \cdots = p_i^{(T)} \quad \text{for } i = 1, \ldots, R,$$

where

$$p_i^{(t)} = \Pr(X_t = i) \quad \text{for } t = 1, \ldots, T.$$

Let $F_i^{(t)}$ denote the marginal cumulative probability and let $C_i^{(t)}$ denote the complementary log-log transform of $F_i^{(t)}$ for $i = 1, \ldots, R-1$; $t = 1, \ldots, T$. Namely, $F_i^{(t)} = \sum_{s=1}^{i} p_s^{(t)}$, and $C_i^{(t)} = \log\left(-\log\left(1 - F_i^{(t)}\right)\right)$. Then the MH model may also be expressed as

$$C_i^{(k)} = C_i^{(1)} \quad \text{for } i = 1, \ldots, R-1;\ k = 2, \ldots, T.$$

Note that

$$C_1^{(t)} < C_2^{(t)} < \cdots < C_{R-1}^{(t)} \quad \text{for } t = 1, \ldots, T.$$

Consider a model defined by

$$C_i^{(k)} = C_i^{(1)} + \log \Delta_{k-1} \quad \text{for } i = 1, \ldots, R-1;\ k = 2, \ldots, T,$$

where the parameter $\{\Delta_{k-1}\}$ is unspecified. We shall refer to this model as the MCL^T model. A special case of this model obtained by putting $\Delta_1 = \cdots = \Delta_{T-1} = 1$ is the MH model.

By putting $\{C_i^{(1)} = \lambda_i\}$, the MCL^T model may be expressed as

$$F_i^{(t)} = 1 - \exp(-\exp(\lambda_i + \log \Delta_{t-1})) \quad \text{for } i = 1, \ldots, R-1;\ t = 1, \ldots, T,$$

where $\Delta_0 = 1$. This model states that the marginal distribution $F_i^{(k)}$ is a location shift of the marginal distribution $F_i^{(1)}$ in terms of above equation for $k = 2, \ldots, T$. Thus, since λ_i is monotone increasing, as the i approaches $R-1$ from 1, $F_i^{(t)}$ may approach 1 more sharply than $F_i^{(1)}$ when $\Delta_{t-1} > 1$, but $F_i^{(t)}$ may approach 1 more slowly than $F_i^{(1)}$ when $\Delta_{t-1} < 1$ for $t = 1, \ldots, T$.

Since the MCL^T model may also be expressed as

$$1 - F_i^{(k)} = \left(1 - F_i^{(1)}\right)^{\Delta_{k-1}} \quad \text{for } i = 1, \ldots, R-1;\ k = 2, \ldots, T,$$

then for k_1 and k_2 $(1 \le k_1 < k_2 \le T)$,

$$\left(1 - F_i^{(k_2)}\right)^{\frac{1}{\Delta_{k_2-1}}} = \left(1 - F_i^{(k_1)}\right)^{\frac{1}{\Delta_{k_1-1}}},$$

thus

$$1 - F_i^{(k_2)} = \left(1 - F_i^{(k_1)}\right)^{\frac{\Delta_{k_2-1}}{\Delta_{k_1-1}}},$$

where $\Delta_0 = 1$ for $i = 1, \ldots, R-1$. Then, this model indicates that the probability that X_k is $i+1$ or above, is equal to the probability that X_1 is $i+1$ or above to the power of Δ_{k-1}, for $i = 1, \ldots, R-1;\ k = 2, \ldots, T$. In other words, this model indicates that the probability that X_{k_2} is $i+1$ or above, is equal to the probability that X_{k_1} is $i+1$ or above to the power of $\frac{\Delta_{k_2-1}}{\Delta_{k_1-1}}$, for $i = 1, \ldots, R-1$. Therefore $\frac{\Delta_{k_2-1}}{\Delta_{k_1-1}} > 1$ is equivalent to $F_i^{(k_2)} > F_i^{(k_1)}$ and $\frac{\Delta_{k_2-1}}{\Delta_{k_1-1}} < 1$ is equivalent to $F_i^{(k_2)} < F_i^{(k_1)}$. As a result, the parameter Δ_{k-1} in the MCL^T model reflects the degree of inhomogeneity between $F_i^{(1)}$ and $F_i^{(k)}$.

2.2. Conditional MCL model

Using the conditional probabilities, the MH model may also be expressed as

$$\begin{aligned} &\Pr\Big(X_k = i \mid (X_1, \ldots, X_T) \neq (s, \ldots, s), s = 1, \ldots, R\Big) \\ &\quad = \Pr\Big(X_1 = i \mid (X_1, \ldots, X_T) \neq (s, \ldots, s), s = 1, \ldots, R\Big), \end{aligned}$$

for $i = 1, \ldots, R$; $k = 2, \ldots, T$; that is

$$p_i^{c(k)} = p_i^{c(1)} \quad \text{for } i = 1, \ldots, R;\ k = 2, \ldots, T,$$

where, for $t = 1, \ldots, T$,

$$p_i^{c(t)} = \Pr\Big(X_t = i \mid (X_1, \ldots, X_T) \neq (s, \ldots, s), s = 1, \ldots, R\Big) = \frac{p_i^{(t)} - p_{ii\ldots i}}{\delta},$$

$$\delta = \Pr\Big((X_1, \ldots, X_T) \neq (s, \ldots, s), s = 1, \ldots, R\Big) = 1 - \sum_{i=1}^{R} p_{ii\ldots i}.$$

Let $F_i^{c(t)}$ denote the conditional marginal cumulative probability of X_t given that $(X_1, \ldots, X_T) \neq (s, \ldots, s), s = 1, \ldots, R$, i.e.,

$$F_i^{c(t)} = \Pr\Big(X_t \leq i \mid (X_1, \ldots, X_T) \neq (s, \ldots, s), s = 1, \ldots, R\Big) = \sum_{l=1}^{i} p_l^{c(t)}$$

for $i = 1, \ldots, R-1;\ t = 1, \ldots, T$. Then the MH model may be further expressed as

$$F_i^{c(k)} = F_i^{c(1)} \quad \text{for } i = 1, \ldots, R-1;\ k = 2, \ldots, T.$$

Consider now a model defined by

$$C_i^{c(k)} = C_i^{c(1)} + \log \Delta_{k-1}^{*} \quad \text{for } i = 1, \ldots, R-1;\ k = 2, \ldots, T,$$

where, for $t = 1, \ldots, T$,

$$C_i^{c(t)} = \log\left(-\log\left(1 - F_i^{c(t)}\right)\right),$$

where the parameter $\{\Delta_{k-1}^{*}\}$ is unspecified. We shall refer to this model as the conditional marginal cumulative complementary log-log (CMCL^T) model. A special case of this model obtained by putting $\Delta_1^{*} = \cdots = \Delta_{T-1}^{*} = 1$ is the MH model.

§3. Decompositions of the MH model

We shall consider two kinds of decompositions of the MH model.

Using the specified monotonic function $g(k)$ in section 1, consider the ME model defined by

$$\sum_{i=1}^{R} g(i) p_i^{(1)} = \cdots = \sum_{i=1}^{R} g(i) p_i^{(T)} \quad (\text{i.e., } \mathrm{E}(g(X_1)) = \cdots = \mathrm{E}(g(X_T))).$$

Using the conditional probabilities, the ME model may also be expressed as

$$\sum_{i=1}^{R} g(i) p_i^{c(1)} = \cdots = \sum_{i=1}^{R} g(i) p_i^{c(T)}.$$

We obtain the following theorem.

Theorem 3.1: *For the R^T table, the MH model holds if and only if both the MCL^T and ME models hold.*

Proof : If the MH model holds, then both the MCL^T and ME models hold. Assuming that both the MCL^T and ME models hold, we shall show that the MH model holds. We have

$$\mathrm{E}(g(X_t)) = g(1) + \sum_{l=1}^{R-1} d_l \left(1 - F_l^{(t)}\right) \quad \text{for } t = 1, \ldots, T,$$

where

$$d_l = g(l+1) - g(l).$$

This is because

$$\begin{aligned}
\mathrm{E}(g(X_t)) &= \sum_{i=1}^{R} g(i) p_i^{(t)} \\
&= \sum_{l=1}^{R-1} g(l) \left(\sum_{i=l}^{R} p_i^{(t)} - \sum_{i=l+1}^{R} p_i^{(t)} \right) + g(R) p_R^{(t)} \\
&= g(1) \sum_{i=1}^{R} p_i^{(t)} + \sum_{l=1}^{R-1} \left(-g(l) \sum_{i=l+1}^{R} p_i^{(t)} + g(l+1) \sum_{i=l+1}^{R} p_i^{(t)} \right) \\
&= g(1) + \sum_{l=1}^{R-1} d_l \left(1 - F_l^{(t)}\right),
\end{aligned}$$

for $t = 1, \ldots, T$.

Then, we have

$$\sum_{l=1}^{R-1} d_l \left(1 - F_l^{(1)}\right) = \sum_{l=1}^{R-1} d_l \left(1 - F_l^{(k)}\right) = \sum_{l=1}^{R-1} d_l \left(1 - F_l^{(1)}\right)^{\Delta_{k-1}},$$

for $k = 2, \ldots, T$, because the ME and MCL^T models hold. Then we obtain $\Delta_{k-1} = 1$ for $k = 2, \ldots, T$, i.e., the MH model holds because $d_l \geq 0$ (or $d_l \leq 0$) for all $l = 1, \ldots, R-1$, with at least one of the $\{d_l\}$ being not equal to zero. The proof is completed.

We also obtain the following theorem.

Theorem 3.2: *For the R^T table, the MH model holds if and only if both the CMCL^T and ME models hold.*

The proof is omitted because it can be obtained in a similar manner to the proof of Theorem 3.1 by replacing $\{F_l^{(1)}\}$ and $\{F_l^{(k)}\}$ with $\{F_l^{c(1)}\}$ and $\{F_l^{c(k)}\}$, respectively.

§4. Goodness-of-fit test

Let $n_{i_1 \dots i_T}$ denote the observed frequency in the $(i_1, \dots, i_T)$ cell of the R^T table with $n = \sum \cdots \sum n_{i_1 \dots i_T}$ and let $m_{i_1 \dots i_T}$ denote the corresponding expected frequency. We assume that $\{n_{i_1 \dots i_T}\}$ have a multinomial distribution. The maximum likelihood estimates (MLEs) of the expected frequencies under each model can be obtained using a Newton-Raphson method to solve the likelihood equations. See Appendix for the likelihood equations under the MCL^T and CMCL^T models. Each of the MH, CMCL^T and ME models do not depend on the probabilities $\{p_{ii \dots i}\}$ on the main diagonal of the table, but the MCL^T model depends on them. Notice that the estimated expected frequencies on the main diagonal cells under the MCL^T model are different from the observed frequencies on the main diagonal.

The likelihood ratio chi-squared statistic for testing the goodness-of-fit of model M is given by

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

where $\hat{m}_{i_1 \dots i_T}$ is the MLE of $m_{i_1 \dots i_T}$ under the model. The numbers of degrees of freedom (df) of statistics for testing the goodness-of-fit of the MH, MCL^T (also CMCL^T), and ME models are $(T-1)(R-1)$, $(T-1)(R-2)$, and $T-1$, respectively. Consider two nested models, say M_1 and M_2, such that if model M_1 holds, then model M_2 holds. For testing the goodness-of-fit of model M_1 assuming that model M_2 holds, the conditional likelihood ratio statistic is given by $G^2(M_1 \mid M_2) = G^2(M_1) - G^2(M_2)$. The number of df for the conditional test is the difference between the numbers of df for the models M_1 and M_2.

§5. Example

Consider the data in Table 1 obtained from the Meteorological Agency in Japan (from Tahata *et al.*, 2008). These are obtained from the daily atmospheric temperatures at Hiroshima, Tokyo, and Sapporo in Japan in 2003, using three levels, (1) low, (2) normal, and (3) high. The variables X_1, X_2, and X_3 mean the temperatures at Hiroshima, Tokyo, and Sapporo, respectively.

Table 2 gives the values of the likelihood ratio chi-square statistic for goodness-of-fit of models applied to these data. We set $g(k) = k$, for $k = 1, 2$, and 3. The MH and ME models fit these data very poorly. However the MCL^3 and CMCL^3 models fit these data well.

Consider the hypothesis that the MH model holds under the assumption that the MCL^3 (CMCL^3) model holds; namely, the hypothesis that $\Delta_1 =$

$\Delta_2 = 1$ ($\Delta_1^* = \Delta_2^* = 1$) under the assumption. Since $G^2(\text{MH}|\text{MCL}^3) = G^2(\text{MH}) - G^2(\text{MCL}^3) = 15.87$ and $G^2(\text{MH}|\text{CMCL}^3) = G^2(\text{MH}) - G^2(\text{CMCL}^3) = 15.78$ with 2 df, we reject these hypotheses at the 0.05 level. These show the rejection of $\Delta_1 = \Delta_2 = 1$ ($\Delta_1^* = \Delta_2^* = 1$) in the MCL^3 (CMCL^3) model. Therefore the MCL^3 (CMCL^3) model is preferable to the MH model for the data.

Under the MCL^3 (CMCL^3) model, the MLEs of $\{\Delta_k\}$ are $\hat{\Delta}_1 = 0.92$ and $\hat{\Delta}_2 = 1.23$ (the MLEs of $\{\Delta_k^*\}$ are $\hat{\Delta}_1^* = 0.89$ and $\hat{\Delta}_2^* = 1.33$). Hence, under the MCL^3 model, the probability that the category for Tokyo is $i+1$ or above, is estimated to be equal to the probability that the category for Hiroshima is $i+1$ or above to the power of 0.92, for $i = 1, 2$, and that the category for Sapporo is $i + 1$ or above, is estimated to be equal to the probability that the category for Hiroshima is $i+1$ or above to the power of 1.23, for $i = 1, 2$. Therefore, the temperature for Hiroshima tends to be stochastically lower than that for Tokyo, but stochastically higher than that for Sapporo.

§6. Concluding Remarks

Under the MCL^T (CMCL^T) model, one marginal distribution is a location shift of the other marginal distribution. When the MH model fits the data poorly, the decompositions of the MH model may be useful for seeing the reason for its poor fit. Indeed, for the data in Table 1, the poor fit of the MH model is caused by the poor fit of the ME model rather than the MCL^T model.

The MLEs of expected frequencies on the main diagonal cell under the CMCL^T model are equal to the observed frequencies, but that of MCL^T model are not. This is because the CMCL^T model is expressed as the function of $\{F_i^{c(k)}\}$, on the other hand, the MCL^T model is expressed as the function of $\{F_i^{(k)}\}$. Thus, if the analyst would be interested in inferring the structure of only off-diagonal probabilities and not the main diagonal probabilities, the decomposition of the MH model into the CMCL^T and ME models may be preferable to that into the MCL^T and ME models. Conversely, if the analyst would be interested in inferring the structure of probabilities including the main diagonal cell, it may be appropriate to use the decomposition of the MH model into the MCL^T and ME models.

The decompositions of the MH model described here should be considered for ordinal categorical data, because each of the decomposed models is not invariant under the same arbitrary permutations of all categories.

Acknowledgments

The authors would like to thank the editor and an anonymous referee for the meaningful comments.

Appendix

For the R^3 table, we give the likelihood equations under each of the MCL^3 and CMCL^3 models.

(a) Case of the MCL^3 model:

To obtain the MLEs of expected frequencies under the MCL^3 model, we must maximize the Lagrangian

$$L = \sum_{i=1}^{R}\sum_{j=1}^{R}\sum_{t=1}^{R} n_{ijt}\log p_{ijt} - \lambda\left(\sum_{i=1}^{R}\sum_{j=1}^{R}\sum_{t=1}^{R} p_{ijt} - 1\right)$$
$$- \sum_{s=1}^{2}\sum_{i=1}^{R-1}\phi_{si}\left(\log\left(1 - F_i^{(s+1)}\right) - \Delta_s \log\left(1 - F_i^{(1)}\right)\right)$$

with respect to $\{p_{ijt}\}$, λ, $\{\phi_{1i}\}$, $\{\phi_{2i}\}$, Δ_1 and Δ_2. Setting the partial derivatives of L equal to zeros, we obtain the equations

$$p_{ijt} = n_{ijt}\left\{n + \sum_{s=1}^{2}\sum_{l=1}^{R-1}\phi_{sl}\left(\frac{F_l^{(s+1)} - I_{s+1}(l)}{1 - F_l^{(s+1)}} - \Delta_s \frac{F_l^{(1)} - I_1(l)}{1 - F_l^{(1)}}\right)\right\}^{-1},$$

where

$$I_1(l) = I(i \le l), \quad I_2(l) = I(j \le l), \quad I_3(l) = I(t \le l),$$

for $i = 1, \ldots, R; j = 1, \ldots, R; t = 1, \ldots, R$,

$$1 - F_i^{(s+1)} = \left(1 - F_i^{(1)}\right)^{\Delta_s},$$

for $i = 1, \ldots, R-1; s = 1, 2$, and

$$\sum_{i=1}^{R-1}\phi_{si}\log\left(1 - F_i^{(1)}\right) = 0,$$

for $s = 1, 2$, where $I(\cdot)$ is the indicator function.

(b) Case of the CMCL^3 model:
We must maximize the Lagrangian

$$L = \sum_{i=1}^{R}\sum_{j=1}^{R}\sum_{t=1}^{R} n_{ijt}\log p_{ijt} - \lambda\left(\sum_{i=1}^{R}\sum_{j=1}^{R}\sum_{t=1}^{R} p_{ijt} - 1\right) - \sum_{s=1}^{2}\sum_{i=1}^{R-1}\phi_{si}\left(\log\left(1-F_i^{c(s+1)}\right) - \Delta_s^*\log\left(1-F_i^{c(1)}\right)\right)$$

with respect to $\{p_{ijt}\}$, λ, $\{\phi_{1i}\}$, $\{\phi_{2i}\}$, Δ_1^* and Δ_2^*. Setting the partial derivatives of L equal to zeros, we obtain the equations

$$p_{ijt} = n_{ijt}\left\{n + \sum_{s=1}^{2}\sum_{l=1}^{R-1}\frac{\phi_{sl}}{\delta}\left(\frac{F_l^{c(s+1)} - I_{s+1}(l)}{1-F_l^{c(s+1)}} - \Delta_s^*\frac{F_l^{c(1)} - I_1(l)}{1-F_l^{c(1)}}\right)\right\}^{-1},$$

where

$$I_1(l) = I(i \le l), \quad I_2(l) = I(j \le l), \quad I_3(l) = I(t \le l),$$

for $i, j, t = 1, \ldots, R;\ (i, j, t) \neq (i, i, i)$,

$$p_{iii} = \frac{n_{iii}}{n},$$

for $i = 1, \ldots, R$,

$$1 - F_i^{c(s+1)} = \left(1 - F_i^{c(1)}\right)^{\Delta_s^*},$$

for $i = 1, \ldots, R-1; s = 1, 2$, and

$$\sum_{i=1}^{R-1}\phi_{si}\log\left(1 - F_i^{c(1)}\right) = 0,$$

for $s = 1, 2$.

References

[1] Agresti, A. (1984). *Analysis of Ordinal Categorical Data*. Wiley, New York.

[2] Agresti, A. (2002). *Categorical Data Analysis*, 2nd edition. Wiley, New York.

[3] Bhapkar, V. P. (1966). A note on the equivalence of two test criteria for hypotheses in categorical data. *Journal of the American Statistical Association* **61**, 228-235.

[4] Bhapkar, V. P. and Darroch, J. N. (1990). Marginal symmetry and quasi symmetry of general order. *Journal of Multivariate Analysis* **34**, 173-184.

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

[6] McCullagh, P. (1977). A logistic model for paired comparisons with ordered categorical data. *Biometrika* **64**, 449-453.

[7] Miyamoto, N., Niibe, K. and 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.

[8] Saigusa, Y., Maruyama, T., Tahata, K. and Tomizawa, S. (2018). Extended marginal homogeneity model based on complementary log-log transform for square tables. *International Journal of Statistics and Probability* **7**, 27-31.

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

[10] Tahata, K., Katakura, S. and Tomizawa, S. (2007). Decompositions of marginal homogeneity model using cumulative logistic models for multi-way contingency tables. *Revstat: Statistical Journal* **5**, 163-176.

[11] Tahata, K., Kobayashi, H. and Tomizawa, S. (2008). Conditional marginal cumulative logistic models and decomposition of marginal homogeneity model for multi-way tables. *Journal of Statistics and Applications* **3**, 239-252.

Table 1

The daily atmospheric temperatures at Hiroshima, Tokyo, and Sapporo in Japan in 2003, using three levels, (1) low, (2) normal, and (3) high (from Tahata *et al.*, 2008). The upper and lower parenthesized values are the MLEs of expected frequencies under the MCL^3 and $CMCL^3$ models, respectively.

Hiroshima	Tokyo	Sapporo		
		(1)	(2)	(3)
(1)	(1)	37	13	3
		(37.26)	(14.14)	(3.05)
		(37.00)	(14.35)	(3.07)
(1)	(2)	21	17	5
		(21.36)	(18.69)	(5.14)
		(21.23)	(18.72)	(5.10)
(1)	(3)	4	4	5
		(4.04)	(4.37)	(5.11)
		(4.05)	(4.41)	(5.11)
(2)	(1)	19	15	5
		(17.51)	(14.82)	(4.65)
		(17.55)	(14.98)	(4.66)
(2)	(2)	20	29	8
		(18.60)	(28.94)	(7.51)
		(18.44)	(29.00)	(7.44)
(2)	(3)	20	20	12
		(18.48)	(19.83)	(11.20)
		(18.46)	(19.96)	(11.18)
(3)	(1)	2	8	4
		(1.96)	(8.45)	(3.96)
		(1.96)	(8.52)	(3.96)
(3)	(2)	8	15	14
		(7.92)	(16.02)	(14.00)
		(7.83)	(15.95)	(13.83)
(3)	(3)	7	21	29
		(6.89)	(22.27)	(28.82)
		(6.86)	(22.36)	(29.00)

Table 2

Likelihood ratio statistic G^2 for models applied to the data in Table 1.

Models	df	G^2
MH	4	16.80*
MCL^3	2	0.93
CMCL^3	2	1.02
ME	2	16.39*

Note: $g(k)$ for the ME model is the equal-interval scores.
* means significant at 0.05 level.

Satoru Shinoda
Department of Information Sciences, Faculty of Science and Technology,
Tokyo University of Science
Noda City, Chiba, 278-8510, Japan
Clinical Data Science Development Headquarters, Taisho Pharmaceutical Co., Ltd.
Toshima-ku, Tokyo, 170-8633, Japan
E-mail: sa-shinoda@taisho.co.jp

Kouji Tahata
Department of Information Sciences, Faculty of Science and Technology,
Tokyo University of Science
Noda City, Chiba, 278-8510, Japan
E-mail: kouji_tahata@is.noda.tus.ac.jp

Kiyotaka Iki
Department of Economics, Nihon University,
Chiyoda-ku, Tokyo, 101-8360, Japan
E-mail: iki.kiyotaka@nihon-u.ac.jp

Sadao Tomizawa
Department of Information Sciences, Faculty of Science and Technology,
Tokyo University of Science
Noda City, Chiba, 278-8510, Japan
E-mail: tomizawa@is.noda.tus.ac.jp