Extended Second-Order Moment Symmetry Model and Decomposition of Symmetry Model for Ordinal Square Contingency Tables

Shuji Ando[1], Ibuki Shisui[1] & Sadao Tomizawa[2]

[1] Department of Information Sciences, Tokyo University of Science, Chiba, Japan

[2] Department of Information Science, Meisei University, Tokyo, Japan

Correspondence: Shuji Ando, Department of Information Sciences, Tokyo University of Science, 2641 Yamazaki, Noda City, Chiba, Japan. E-mail: shuji.ando@rs.tus.ac.jp

Received: February 21, 2024 Accepted: May 29, 2024 Online Published: August 31, 2024

doi:10.5539/ijsp.v13n3p16 URL: https://doi.org/10.5539/ijsp.v13n3p16

Abstract

This study proposes the extended second-order moment symmetry model in square contingency tables with same row and column ordinal classifications. The proposed model extends the previous second-order moment symmetry model by incorporating an asymmetry parameter to capture complex structures. The proposed and previous second-order moment symmetry models represent the asymmetry and symmetry structure of the covariance matrix, respectively. The previous study showed the decomposition of the symmetry model using the second-order moment symmetry model. This study shows the decomposition of the second-order moment symmetry model using the proposed model and the decomposition of the symmetry model applying its decomposition. The usefulness of proposed model is illustrated with examples of two real world datasets (cross-classified education of parents and cross-classified social group of fathers and their daughters).

Keywords: asymmetry, covariance, cross-classified, moment, ordinal category

1. Introduction

In two-way contingency tables analysis, we are often interested in whether or not the independence model holds for the presented data. The independence model has multiple expressions, and the same restrictions can be viewed from different expressions and perspectives. In square contingency tables with same row and column classifications, the independence model usually does not hold because there is a tendency for the observed frequencies to concentrate in the cells of the main diagonal. Hence, we are interested in whether or not models such as the symmetry (S) model (Bowker, 1948), the marginal homogeneity (MH) model (Stuart, 1955) and the quasi-symmetry (QS) model (Caussinus, 1965) introduced below hold. Recently, Yoshimoto, Tahata, Iki and Tomizawa (2019) introduced new expressions of the symmetry model and the marginal homogeneity model. We first introduce basic models for analyzing square contingency tables, and then describe the second-order moment symmetry (SMS) model.

Consider $r \times r$ square contingency tables. Let denote the row and column variables as X_1 and X_2, respectively, and denote the (i, j)th $(i, j = 1, 2, \ldots, r)$ cell probability as $\Pr(X_1 = i, X_2 = j) = p_{ij}$. Bowker (1948) considered the S model defined by

$$p_{ij} = p_{ji} \quad (i = 1, 2, \ldots, r, j = 1, 2, \ldots, r).$$

This model indicates the symmetry of cell probabilities (i.e., the probability that an observation will fall in the cell of row category i and column category j is equal to the probability that an observation falls in the cell of row category j and column category i). The S model often does not fit for the presented data well because it has strong restrictions.

The MH model (Stuart, 1955) was created to represent the symmetry of marginal distributions, rather than the symmetry of cell probabilities. The MH model has weaker restrictions than the S model. The MH model is defined by

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

where $p_{i\cdot} = \sum_{j=1}^{r} p_{ij}$ and $p_{\cdot i} = \sum_{j=1}^{r} p_{ji}$. The MH model indicates that the row marginal distribution is identical to the column marginal distribution. If the S model holds, then the MH model always holds. However, the converse is not always true. Caussinus (1965) considered what additional restrictions should be added to the MH model to be equivalent to the S model, and detected that the additional restrictions are the QS model.

Caussinus (1965) presented the QS model defined by

$$p_{ij} = \alpha_i \beta_j \psi_{ij} \quad (i = 1, 2, \ldots, r, j = 1, 2, \ldots, r),$$

where $\psi_{ij} = \psi_{ji}$. The interpret of the QS model may be difficult compared to the S and MH models, although the QS model can be understood as the symmetry of odds ratios. In fact, when we define the odds ratio between ith row and j ($> i$)th row, and sth column and t ($> s$)th column as

$$\theta_{ij;st} = \frac{p_{is} p_{jt}}{p_{js} p_{it}},$$

then the QS model can be expressed as $\theta_{ij;st} = \theta_{st;ij}$ $(i < j; s < t)$. We call a decomposition of model that restrictions of a single model are identical to restrictions combining multiple models having weaker restrictions than the single model. From the decomposition of the S model considered by Caussinus (1965), we can investigate which the cause, that the S model does not fit for the presented data well, is either of the MH model or the QS model (or both).

Yoshimoto et al. (2019) provided an alternative expression of the S model using the following definition. Let denote random variable vectors as

$$\boldsymbol{X}^{(1)} = (X_1^{(1)}, X_2^{(1)}, \ldots, X_r^{(1)})^\top \quad \text{and} \quad \boldsymbol{X}^{(2)} = (X_1^{(2)}, X_2^{(2)}, \ldots, X_r^{(2)})^\top. \tag{1}$$

The $X_i^{(1)}$ $(i = 1, 2, \ldots, r)$ takes 0 or 1, $X_j^{(2)}$ $(j = 1, 2, \ldots, r)$ takes 0 or 1 where $\sum_{i=1}^r X_i^{(1)} = \sum_{j=1}^r X_j^{(2)} = 1$, the $\boldsymbol{\delta}_s$ is the $r \times 1$ vector where the sth component is 1 and others are 0, and the symbol $\top$ is the transpose of vector (or matrix). Let denote the probability distribution of $(\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)})$ by

$$\begin{aligned}\Pr(\boldsymbol{X}^{(1)} = \boldsymbol{\delta}_i, \boldsymbol{X}^{(2)} = \boldsymbol{\delta}_j) &= \Pr(X_1 = i, X_2 = j) \\ &= p_{ij} \quad (i = 1, 2, \ldots, r, j = 1, 2, \ldots, r).\end{aligned}$$

Under the definition (1), the S and MH models can be expressed as

$$E(X_i^{(1)} X_j^{(2)}) = E(X_j^{(1)} X_i^{(2)}) \quad (i \neq j) \quad \text{and} \quad \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2,$$

respectively, where $\boldsymbol{\mu}_1 = E(\boldsymbol{X}^{(1)}) = (p_{1\cdot}, \ldots, p_{r\cdot})^\top$ and $\boldsymbol{\mu}_2 = E(\boldsymbol{X}^{(2)}) = (p_{\cdot 1}, \ldots, p_{\cdot r})^\top$. The S model indicates the symmetry of the second-order moments with respect to zero, and the MH model implies the symmetry of the first-order moments regarding zero. Therefore, additional restrictions, that should be added to the MH model to be equivalent to the S model, are the symmetry of the second-order moments with respect to means (i.e., covariance matrix of $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$).

Yoshimoto et al. (2019) proposed the SMS model defined by

$$\mathrm{Cov}(X_i^{(1)}, X_j^{(2)}) = \mathrm{Cov}(X_j^{(1)}, X_i^{(2)}) \quad (i < j),$$

where $\mathrm{Cov}(X_i^{(1)}, X_j^{(2)}) = E[(X_i^{(1)} - E[X_i^{(1)}])(X_j^{(2)} - E[X_j^{(2)}])] = p_{ij} - p_{i\cdot} p_{\cdot j}$. The SMS model can also be defined as

$$\mathrm{Cov}(X_i^{(1)}, X_j^{(2)}) = \mathrm{Cov}(X_j^{(1)}, X_i^{(2)}) \quad ((i, j) \in A_h),$$

where $A_h = \{(i, j) \mid i, j = 1, 2, \ldots, r, i < j\} \setminus \{\{i = h\} \cup \{j = h\}\}$ and the h is one of the categories from 1 to r.

Consider the data in Table 1, which is from Mullins and Sites (1984). Table 1 is the ordinal square contingency table data that classified mother's and father's education for a sample of eminent black Americans. Categories I to IV mean 8th grade or less, part high school, high school and college, respectively.

Although we will provide detailed results in Section 4, the S, MH and SMS models do not fit for the data in Table 1 well. We realized that a model having the asymmetry of covariance matrix of $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$ can improve the fitting. Consider the Table 2, which shows estimates of covariance corresponding to each combination of $X_i^{(1)}$ and $X_j^{(2)}$ for the data in Table 1. It must be noted that estimate of covariance between $X_i^{(1)}$ and $X_j^{(2)}$ is $\tilde{p}_{ij} - \tilde{p}_{i\cdot}\tilde{p}_{\cdot j}$, where $\tilde{p}_{ij} = n_{ij}/N$, n_{ij} is the observed frequency in the (i, j)th cell, $\tilde{p}_{i\cdot} = \sum_{j=1}^r \tilde{p}_{ij}$, $\tilde{p}_{\cdot i} = \sum_{j=1}^r \tilde{p}_{ji}$, and the N is sample size (i.e., $N = \sum\sum n_{ij}$).

From the result of Table 2, we found that (i) in the case where the education level I is specified as the category h, the ratios of $\mathrm{Cov}(X_i^{(1)}, X_j^{(2)})$ and $\mathrm{Cov}(X_j^{(1)}, X_i^{(2)})$ are 5.53, 2,24 and 2,75 from top to bottom in Table 2, (ii) in the case where the education level II is specified as the category h, the ratios are 6.06, 0.61 and 2.75, (iii) in the case where the education level III is specified as the category h, the ratios are 3.13, 0.61 and 2.23, and (iv) in the case where the education level IV is specified as the category h, the ratios are 3.13, 6.06 and 5.53. Therefore, a model having the asymmetry of covariance

Table 1. Mother's and father's education for a sample of eminent black Americans

Mother's	Father's education				
education	I	II	III	IV	Total
I	81	3	9	11	104
II	14	8	9	6	37
III	43	7	43	18	111
IV	21	6	24	87	138
Total	159	24	85	122	390

Source: Mullins and Sites (1984).

Table 2. Estimates of covariance corresponding to each combination of $X_i^{(1)}$ and $X_j^{(2)}$ $(i < j)$ in Table 1

(i, j)	$\mathrm{Cov}(X_i^{(1)}, X_j^{(2)})$	$\mathrm{Cov}(X_j^{(1)}, X_i^{(2)})$	$\dfrac{\mathrm{Cov}(X_i^{(1)}, X_j^{(2)})}{\mathrm{Cov}(X_j^{(1)}, X_i^{(2)})}$
$(1,2)$	-0.0087	-0.0028	3.1348
$(1,3)$	-0.0350	-0.0058	6.0637
$(1,4)$	-0.0552	-0.0904	0.6107
$(2,3)$	0.0024	0.0004	5.5303
$(2,4)$	-0.0143	-0.0064	2.2366
$(3,4)$	-0.0429	-0.0156	2.7519

matrix of $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$, in the case where the education level IV (or I) is specified as the category h, may be improved the fitting.

The rest of this paper is organized as follows. Section 2 gives the proposed model and the decomposition of the S model using the proposed model. Section 3 gives the goodness-of-fit test for the proposed model. Section 4 gives some examples to illustrate the usefulness of the proposed model. Finally, Section 5 concludes this paper.

2. Proposed Model and Decomposition of Symmetry Model

For $r \times r$ square contingency tables, it can generally be written as

$$E\begin{pmatrix}\boldsymbol{X}^{(1)}\\ \boldsymbol{X}^{(2)}\end{pmatrix} = \begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{pmatrix} \quad \text{and} \quad V\begin{pmatrix}\boldsymbol{X}^{(1)}\\ \boldsymbol{X}^{(2)}\end{pmatrix} = \begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix},$$

where

$$\begin{aligned}\boldsymbol{\Sigma}_{11} &= \boldsymbol{P}_{11} - \boldsymbol{\mu}_1\boldsymbol{\mu}_1^\top, \quad \boldsymbol{\Sigma}_{12} = \boldsymbol{P}_{12} - \boldsymbol{\mu}_1\boldsymbol{\mu}_2^\top,\\ \boldsymbol{\Sigma}_{21} &= \boldsymbol{P}_{21} - \boldsymbol{\mu}_2\boldsymbol{\mu}_1^\top \ (= \boldsymbol{\Sigma}_{12}^\top) \quad \text{and} \quad \boldsymbol{\Sigma}_{22} = \boldsymbol{P}_{22} - \boldsymbol{\mu}_2\boldsymbol{\mu}_2^\top,\end{aligned}$$

with $\boldsymbol{P}_{11}$ is the $r \times r$ diagonal matrix where the sth diagonal element is $p_{s\cdot}$ $(s = 1, \ldots, r)$, $\boldsymbol{P}_{12}$ is the $r \times r$ matrix where the (s, t)th element is p_{st} $(s, t = 1, 2, \ldots, r)$, $\boldsymbol{P}_{21}$ is the $r \times r$ matrix where the (s, t)th element is p_{ts} $(s, t = 1, 2, \ldots, r)$, and $\boldsymbol{P}_{22}$ is the $r \times r$ diagonal matrix where the tth diagonal element is $p_{\cdot t}$ $(t = 1, 2, \ldots, r)$. The SMS model can be expressed as

$$\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{12}^\top.$$

Therefore, the SMS model indicates the symmetric structure of the covariance matrix of $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$. Now we are interested in proposing the asymmetric structure of the covariance matrix of $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$.

Assuming that $\mathrm{Cov}(X_i^{(1)}, X_j^{(2)}) \neq 0$ $(i, j = 1, 2, \ldots, r)$, we propose the extended second-order moment symmetry (ESMS) model, which is expressed as

$$\mathrm{Cov}(X_i^{(1)}, X_j^{(2)}) = \Delta \mathrm{Cov}(X_j^{(1)}, X_i^{(2)}) \quad ((i, j) \in A_h),$$

the parameter Δ represents the degree of asymmetry structure of $\boldsymbol{\Sigma}_{12}$. Under the ESMS model, $\boldsymbol{\Sigma}_{12}$ can be specifically expressed as

$$\mathbf{\Sigma}_{12} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1r} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{r1} & \sigma_{r2} & \cdots & \sigma_{rr} \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_{11} & \Delta\sigma_{21} & \cdots & \Delta\sigma_{h-1,1} & \sigma_{1h} & \Delta\sigma_{h+1,1} & \cdots & \Delta\sigma_{r-1,1} & \Delta\sigma_{r1} \\ \sigma_{21} & \sigma_{22} & \cdots & \Delta\sigma_{h-1,2} & \sigma_{2h} & \Delta\sigma_{h+1,2} & \cdots & \Delta\sigma_{r-1,2} & \Delta\sigma_{r2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{h1} & \sigma_{h2} & \cdots & \sigma_{h,h-1} & \sigma_{hh} & \Delta\sigma_{h+1,h} & \cdots & \Delta\sigma_{r-1,h} & \Delta\sigma_{rh} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{r-1,1} & \sigma_{r-1,2} & \cdots & \sigma_{r-1,h-1} & \sigma_{r-1,h} & \sigma_{r-1,h+1} & \cdots & \sigma_{r-1,r-1} & \Delta\sigma_{r,r-1} \\ \sigma_{r1} & \sigma_{r2} & \cdots & \sigma_{r,h-1} & \sigma_{rh} & \sigma_{r,h+1} & \cdots & \sigma_{r,r-1} & \sigma_{rr} \end{pmatrix},$$

where $\sigma_{ij} = \mathrm{Cov}(X_i^{(1)}, X_j^{(2)}) = p_{ij} - p_{i\cdot}p_{\cdot j}$ $(i, j = 1, 2, \ldots, r)$.

If the S model holds, then both the ESMS and MH models always hold. However, the converse is not always true. To consider what additional restrictions should be added to both the ESMS and MH models to be equivalent to the S model, we propose the global second-order moment symmetry (GSMS) model, which is expressed as

$$\sum\sum_{i<j} \mathrm{Cov}(X_i^{(1)}, X_j^{(2)}) = \sum\sum_{i<j} \mathrm{Cov}(X_j^{(1)}, X_i^{(2)}).$$

This model can also be expressed as

$$\sum\sum_{i<j} p_{ij} - \sum\sum_{i<j} p_{ji} = \sum\sum_{i<j} p_{i\cdot}p_{\cdot j} - \sum\sum_{i<j} p_{j\cdot}p_{\cdot i}. \tag{2}$$

We introduce the relationship between the left side of equation (2) and the global symmetry (GS) model (Read, 1977), and the relationship between the right side of equation (2) and the marginal ridits (MR) model (Iki, Tahata and Tomizawa, 2012). Read (1977) proposed the GS model defined as

$$\sum\sum_{i<j} p_{ij} = \sum\sum_{i<j} p_{ji}.$$

The GS model indicates that the sum of probabilities in upper triangular portion except the main diagonal cells of the square contingency table is equal to the sum of probabilities in lower triangular portion except the main diagonal cells. Studies related to the GS model include, for example, Tomizawa and Saitoh (1998) and Ando (2020).

According to Iki et al. (2012), the MR model is expressed as

$$\sum\sum_{i<j} p_{i\cdot}p_{\cdot j} = \sum\sum_{i<j} p_{j\cdot}p_{\cdot i}.$$

By using $\tau = \sum\sum_{i<j} p_{i\cdot}p_{\cdot j} - \sum\sum_{i<j} p_{j\cdot}p_{\cdot i}$, we can compare the marginal distributions from the perspective of mean rigits (Agresti, 1983; Svensson, 1994; Svensson, 1997; Svensson, 1998; Agresti and Kateri, 2017; Lu and Ding, 2020). In addition, the MR model can be expressed as

$$P(Y_1 > Y_2) = P(Y_1 < Y_2),$$

where Y_1 is selected at random from the row marginal distribution and Y_2 is selected independently at random from the column marginal distribution. Therefore, the GSMS model indicates that the degree of departure from the GS model is equal to the degree of departure from the MR model in the sense of difference.

From the definition of the SMS model, it is clear that the GSMS can further be expressed as

$$\sum\sum_{(i,j)\in A_h} \mathrm{Cov}(X_i^{(1)}, X_j^{(2)}) = \sum\sum_{(i,j)\in A_h} \mathrm{Cov}(X_j^{(1)}, X_i^{(2)}).$$

Now, we introduce a decomposition of the SMS model using the ESMS model and a decomposition of the S model using the ESMS model.

Lemma 1. *For an $r\times r$ square contingency table, assuming that $\sum\sum_{(i,j)\in A_h}(p_{ij}-p_{i\cdot}p_{\cdot j}) \neq 0$ and $\sum\sum_{(i,j)\in A_h}(p_{ji}-p_{j\cdot}p_{\cdot i}) \neq 0$, the SMS model holds if and only if both the ESMS and GSMS models hold.*

Proof. If the SMS model holds, then both the ESMS and GSMS models hold. Conversely, assume that both the ESMS and GSMS models hold. As the ESMS model holds, the following equation holds:

$$\sum\sum_{(i,j)\in A_h}\mathrm{Cov}(X_i^{(1)}, X_j^{(2)}) = \Delta\sum\sum_{(i,j)\in A_h}\mathrm{Cov}(X_j^{(1)}, X_i^{(2)})$$
$$\Longleftrightarrow \sum\sum_{(i,j)\in A_h}(p_{ij} - p_{i\cdot}p_{\cdot j}) = \Delta\sum\sum_{(i,j)\in A_h}(p_{ji} - p_{j\cdot}p_{\cdot i}). \quad (3)$$

As the GSMS model holds, the equation (3) can be expressed as

$$\sum\sum_{(i,j)\in A_h}(p_{ij} - p_{i\cdot}p_{\cdot j}) = \Delta\sum\sum_{(i,j)\in A_h}(p_{ji} - p_{j\cdot}p_{\cdot i})$$
$$\Longleftrightarrow (1-\Delta)\sum\sum_{(i,j)\in A_h}(p_{ij} - p_{i\cdot}p_{\cdot j}) = 0.$$

Therefore, we obtain $\Delta = 1$ because $\sum\sum_{(i,j)\in A_h}(p_{ij} - p_{i\cdot}p_{\cdot j}) \neq 0$. From the above results, if both the ESMS and GSMS models hold, then the SMS model also holds. The proof is complete. □

In square contingency tables, the independence model usually does not hold. This is because, there is a tendency for the observed frequencies to concentrate in the cells of the main diagonal. As, in square contingency tables, it is natural that all of $\mathrm{Cov}(X_i^{(1)}, X_j^{(2)})$ for $(i, j) \in A_h$ are positive or negative, the assumption of Lemma 1 is valid. Yoshimoto et al. (2019) gave the decomposition of the S model using the SMS and MH models. We obtain the following theorem by combining this decomposition and Lemma 1.

Theorem 1. *For an $r\times r$ square contingency table, assuming that $\sum\sum_{(i,j)\in A_h}(p_{ij}-p_{i\cdot}p_{\cdot j}) \neq 0$ and $\sum\sum_{(i,j)\in A_h}(p_{ji}-p_{j\cdot}p_{\cdot i}) \neq 0$, the S model holds if and only if all of the MH, ESMS, and GSMS models hold.*

Proof. If the S model holds, then all of the MH, ESMS and GSMS models hold. Conversely, assume that all of the MH, ESMS and GSMS models hold. As both the ESMS and GSMS models hold, then the SMS model holds from the Lemma 1. In addition, as both the SMS and MH models hold, then the S model holds from the decomposition given by Yoshimoto et al. (2019). The proof is complete. □

The relationships among models are given in Figure 1.

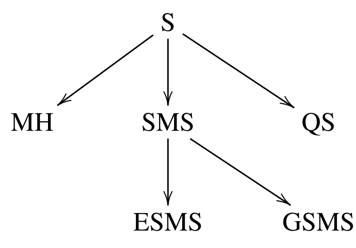

Figure 1. Relationships among models (A → B indicates that model A implies model B)

3. Goodness-of-fit Test

Each model can be tested for the goodness-of-fit using a test statistic such as the likelihood ratio chi-squared statistic (denoted by G^2) with the corresponding degrees of freedom (df). The G^2 is expressed as

$$G^2(\mathrm{H}) = 2\sum_{i=1}^{r}\sum_{j=1}^{r} n_{ij}\log\left(\frac{n_{ij}}{\hat{m}_{ij}}\right),$$

where $\hat{m}_{ij}$ is the maximum likelihood estimate (MLE) of expected frequency m_{ij} under model H.

To obtain $\hat{m}_{ij}$ under the ESMS model, we maximize the following Lagrangian regarding $\{p_{ij}\}$, $\{\phi\}$, $\{\lambda_{ij}\}$ and Δ:

$$L = \sum_{i=1}^{r}\sum_{j=1}^{r} n_{ij}\log p_{ij} - \phi\left(\sum_{i=1}^{r}\sum_{j=1}^{r} p_{ij} - 1\right) - \sum\sum_{(i,j)\in A_h} \lambda_{ij}\left\{p_{ij} - p_{i\cdot}p_{\cdot j} - \Delta\left(p_{ji} - p_{j\cdot}p_{\cdot i}\right)\right\}.$$

Using the Newton-Raphson method, we can solve the equations with respect to $\{p_{ij}\}$, $\{\phi\}$, $\{\lambda_{ij}\}$ and Δ, and then obtain the $\hat{m}_{ij}$ under the ESMS model.

To obtain $\hat{m}_{ij}$ under the GSMS model, we maximize the following Lagrangian regarding $\{p_{ij}\}$, $\{\phi\}$ and $\{\lambda\}$:

$$L = \sum_{i=1}^{r}\sum_{j=1}^{r} n_{ij}\log p_{ij} - \phi\left(\sum_{i=1}^{r}\sum_{j=1}^{r} p_{ij} - 1\right) - \lambda\sum\sum_{(i,j)\in A_h} \left\{p_{ij} - p_{i\cdot}p_{\cdot j} - \left(p_{ji} - p_{j\cdot}p_{\cdot i}\right)\right\}.$$

Using the Newton-Raphson method, we can solve the equations with respect to$\{p_{ij}\}$, $\{\phi\}$ and $\{\lambda\}$, and then obtain the $\hat{m}_{ij}$ under the GSMS model.

The df for each model are listed in Table 3.

Table 3. Degrees of freedom of each model.

Models	S	MH	SMS	ESMS	GSMS
df	$\frac{r(r-1)}{2}$	$r-1$	$\frac{(r-1)(r-2)}{2}$	$\frac{r(r-3)}{2}$	1

4. Examples

For applying the ESMS model to data, it is natural to specify $h = 1$, $h = r$, or a reference category for ordinal square contingency tables. In this section, we provide examples for the cases of $h = 1$ and $h = r$.

4.1 Example 1

Consider the data in Table 1. In Table 4, upper parenthesized values are MLEs of m_{ij} under the ESMS model in the case where the education level Ⅳ is specified as the category h, and lower parenthesized values are MLEs of m_{ij} under the ESMS model in the case where the education level I is specified as the category h.

Table 4. MLEs of m_{ij} under the ESMS model applied to Table 1.

Mother's education	Father's education I	Ⅱ	Ⅲ	Ⅳ	Total
I	81	3	9	11	104
	(80.926)	(2.999)	(9.016)	(11.001)	
	(81.062)	(2.807)	(9.146)	(11.015)	
Ⅱ	14	8	9	6	37
	(14.449)	(7.893)	(8.823)	(5.860)	
	(14.332)	(7.646)	(8.984)	(5.883)	
Ⅲ	43	7	43	18	111
	(42.644)	(6.946)	(43.285)	(18.142)	
	(42.596)	(7.179)	(43.246)	(18.064)	
Ⅳ	21	6	24	87	138
	(20.980)	(6.062)	(23.971)	(87.004)	
	(21.018)	(6.277)	(23.620)	(87.126)	
Total	159	24	85	122	390

Upper parenthesized values are MLEs of m_{ij} under the ESMS model in the case where the education level Ⅳ is specified as the category h.
Lower parenthesized values are MLEs of m_{ij} under the ESMS model in the case where the education level I is specified as the category h.

Table 5 shows the values of G^2 for each model applied to Table 1. From Table 5, the S, MH, SMS and GSMS models fit the data poorly, whereas the ESMS model fits the data well. From Lemma 1, the poor fit of the SMS model is due to the lack of structure of the GSMS model rather than the ESMS model. Moreover, from Theorem 1, the poor fit of the S model is due to the lack of structures of the MH and GSMS models rather than the ESMS model. Under the ESMS model in the case where the education level Ⅳ is specified as the category h, the MLE of Δ is 5.222. This means that the strength of association of $X_i^{(1)}$ and $X_j^{(2)}$ for $(i, j) \in A_h$ are 5.222 times larger than that of $X_j^{(1)}$ and $X_i^{(2)}$ when using the covariance. Under the ESMS model in the case where the education level I is specified as the category h, the MLE of Δ is 2.585.

Table 5. Values of G^2 for each model applied to Table 1

Models	df	G^2
S	6	36.183*
MH	3	32.145*
SMS	3	11.384*
ESMS[1]	2	0.030
ESMS[2]	2	0.070
GSMS	1	6.069*

∗ means significance at the 0.05 level.
[1] means the ESMS model in the case where the education level Ⅳ is specified as the category h.
[2] means the ESMS in the case where the education level I is specified as the category h.

4.2 Example 2

Consider the data in Table 6, which is from Andersen (1997, p. 207).

Table 6. The social groups of the married women between the age of 40 and 59 in the Danish Welfare Study cross-classified with the social group of their fathers

Father's	Daughter's social group				
social group	I-Ⅱ	Ⅲ	Ⅳ	V	Total
I-Ⅱ	12	17	22	3	54
	(12.905)	(17.370)	(21.168)	(3.048)	
	(11.796)	(16.542)	(24.078)	(2.841)	
Ⅲ	8	33	85	95	221
	(8.753)	(32.639)	(85.044)	(94.383)	
	(8.770)	(35.769)	(77.158)	(99.392)	
Ⅳ	11	26	72	87	196
	(9.456)	(26.367)	(72.727)	(87.540)	
	(10.157)	(24.539)	(73.513)	(88.080)	
V	2	18	50	111	181
	(2.107)	(17.857)	(50.018)	(110.617)	
	(1.976)	(17.617)	(54.123)	(105.650)	
Total	33	94	229	296	652

Source: Andersen (1997, p. 207).
Upper parenthesized values are MLEs of m_{ij} under the ESMS model in the case where the social group V is specified as the category h.
Lower parenthesized values are MLEs of m_{ij} under the ESMS model in the case where the social group I-Ⅱ is specified as the category h.

Table 6 is the ordinal square contingency table data that classified social groups of the married women between the age of 40 and 59, and social groups of their fathers in Denmark. Social group I means academics and main executives in the private and public sector. Group Ⅱ means second level executives in the private and public sector. Group Ⅲ means foremen, heads of sections etc. with less than 5 employees below them. Group Ⅳ means white collar workers and blue collar workers with special training. Group V means blue collar workers.

In Table 6, upper parenthesized values are MLEs of m_{ij} under the ESMS model in the case where the social group V is specified as the category h, and lower parenthesized values are MLEs of m_{ij} under the ESMS model in the case where the social group I-Ⅱ is specified as the category h. Table 7 shows the values of G^2 for each model applied to Table 6. From Table 7, we show that the S, MH, SMS and GSMS models fit the data poorly, whereas the ESMS model fits the data well. From Lemma 1, the poor fit of the SMS model is due to the lack of structure of the GSMS model rather than the ESMS model. Moreover, from Theorem 1, the poor fit of the S model is due to the lack of structures of the MH and GSMS models rather than the ESMS model. Under the ESMS model in the case where the social group V is specified as the category h, the MLE of Δ is -3.800. This means that the strength of association of $X_i^{(1)}$ and $X_j^{(2)}$ for $(i, j) \in A_h$ are 3.800 times stronger than that of $X_j^{(1)}$ and $X_i^{(2)}$ when using the covariance. Under the ESMS model in the case where the social group I-II is specified as the category h, the MLE of Δ is 0.116.

Table 7. Values of G^2 for each model applied to Table 6

Models	df	G^2
S	6	107.957*
MH	3	101.284*
SMS	3	22.394*
ESMS[1]	2	0.444
ESMS[2]	2	2.263
GSMS	1	10.701*

* means significance at the 0.05 level.
[1] means the ESMS model in the case where the social group V is specified as the category h.
[2] means the ESM in the case where the social group I-II is specified as the category h.

5. Concluding Remarks

This study proposed both the ESMS and GSMS models and decomposed the S model into the MH, ESMS and GSMS models. Such as shown in Examples, we could determine the reason, which the poor fit of the S model for the presented data, using Theorem 1. The ESMS model indicated that the data had the asymmetry of covariance matrix of $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$. The ESMS model fits for the presented data well if the covariance matrix has the asymmetric structure, that is, the Δ of the ESMS model is away from 1. Relationships between row and column categories are also showed using $X_i^{(1)}$, $X_j^{(2)}$, $X_j^{(1)}$ and $X_i^{(2)}$.

In addition, the GSMS model has a interesting feature. We introduced the relationship between the left side of equation (2) and the GS model, and the relationship between the right side of equation (2) and the MR model. The GSMS model indicates that the degree of departure from the GS model is equal to the degree of departure from the MR model in the sense of difference.

Which category is appropriate for category h under various data and research themes, the practicality of excluding category h from the ESMS model's restrictions, and more detailed examinations of the GSMS model will be subjects for future studies.

Acknowledgments

The authors are grateful to the anonymous referee and the editor for their valuable suggestions on improving this article.

Authors contributions

Dr. Ando and Mr. Shisui drafted the manuscript and Prof. Tomizawa revised it. All authors read and approved the final manuscript.

Funding

The authors have solely funded the research by themselves.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Agresti, A. (1983). Testing marginal homogeneity for ordinal categorical variables. *Biometrics, 39*(2), 505–510. https://doi.org/10.2307/2531022

Agresti, A., & Kateri, M. (2017). Ordinal probability effect measures for group comparisons in multinomial cumulative link models. *Biometrics, 73*(1), 214–219. https://doi.org/10.1111/biom.12565

Andersen, E. B. (1997). *Introduction to the statistical analysis of categorical data.* Springer, Germany.

Ando, S. (2020). Directional index for measuring global symmetry in square tables. *Journal of Statistical Theory and Practice, 14*, 50. https://doi.org/10.1007/s42519-020-00117-4

Bowker, A. H. (1948). A test for symmetry in contingency tables. *Journal of the American Statistical Association, 43*(244), 572–574.

Caussinus, H. (1965). Contribution à l'analyse statistique des tableaux de corrélation. *Annales de la Faculté des Sciences de l'Université de Toulouse, 29*, 77–183.

Iki, K., Tahata, K., & Tomizawa, S. (2012). On quasi-symmetry based on ridit for analysis of square contingency tables. *Journal of Statistical Theory and Applications, 11*(1), 9–22.

Lu, J., Zhang, Y., & Ding, P. (2020). Sharp bounds on the relative treatment effect for ordinal outcomes. *Biometrics, 76*(2), 664–669. https://doi.org/10.1111/biom.13148.

Mullins, E. I., & Sites, P. (1984). The origins of contemporary eminent black americans: a three-generation analysis of social origin. *American Sociological Review, 49*(5), 672–685. https://doi.org/10.2307/2095424

Read, C. B. (1977). Partitioning chi-square in contingency tables: a teaching approach. *Communications in Statistics-Theory and Methods, 6*(6), 553–562. https://doi.org/10.1080/03610927708827513

Stuart, A. (1955). A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika, 42*(3/4), 412–416. https://doi.org/10.2307/2333387

Svensson, E. (1997). A coefficient of agreement adjusted for bias in paired ordered categorical data. *Biometrical Journal, 39*(6), 643–657. https://doi.org/10.1002/bimj.4710390602

Svensson, E. (1998). Ordinal invariant measures for individual and group changes in ordered categorical data. *Statistics in Medicine, 17*(24), 2923–2936.

Svensson, E., & Holm, S. (1994). Separation of systematic and random differences in ordinal rating scales. *Statistics in Medicine, 13*, 2437–2453. https://doi.org/10.1002/sim.4780132308

Tomizawa, S., & Saitoh, K. (1998). Generalized measure of departure from global symmetry for square contingency tables with ordered categories. *Journal of the Korean Statistical Society, 27*(3), 289–303.

Yoshimoto, T., Tahata, K., Iki, K., & Tomizawa, S. (2019). Moment symmetry models and decompositions of symmetry for multi-way contingency tables. *Calcutta Statistical Association Bulletin*, 71(2), 83–98. https://doi.org/10.1177/0008068319880443