symmetry

MDPI

Article

Quasi Association Models for Square Contingency Tables with Ordinal Categories

Kengo Fujisawa [1,*] and Kouji Tahata [2]

Faculty of Science and Technology, Tokyo University of Science, Chiba 278-8510, Japan; kouji_tahata@is.noda.tus.ac.jp
* Correspondence: kfujisawa@rs.tus.ac.jp

Abstract: The analysis of contingency tables focuses on a statistical model instead of independence when the independence between row and column variables does not hold. Many association models have been proposed to indicate the structure of odds ratios. Additionally, symmetry and asymmetry models have been proposed to analyze the cell probabilities of square contingency tables with symmetric or asymmetric structures. This paper proposes an asymmetry plus association model for square contingency tables with ordinal categories and partitioning of the test statistic for goodness-of-fit using our proposed model.

Keywords: association model; asymmetry model; square contingency table

check for updates

Citation: Fujisawa, K.; Tahata, K. Quasi Association Models for Square Contingency Tables with Ordinal Categories. *Symmetry* **2022**, *14*, 805. https://doi.org/10.3390/sym14040805

Academic Editor: Alice Miller

Received: 25 February 2022
Accepted: 11 April 2022
Published: 12 April 2022

1. Introduction

A categorical variable distinguishes a set of categories. It is employed in diverse fields such as social sciences, medical sciences, engineering, and education. Here, we consider a categorical variable with r categories and another one with c categories. The outcome for two variables has rc possible combinations, which can be denoted by a rectangular table with r rows and c columns, where the cells illustrate the rc possible outcomes. This is called a contingency table (for more details, see [1,2]). A contingency table illustrates the joint frequencies by combination of two categorical variables. When analyzing a contingency table, only the observed frequencies are seen, but the true distribution is unknown. One of the aims of analyzing a contingency table is to estimate an unknown probability distribution from the observed frequencies. The confidence level of the estimated distribution is higher when fewer parameters are used to describe the data. Sometimes, we need to consider a parsimonious model. Traditionally, a contingency table is used to evaluate whether classifications are associated. That is, the analysis determines whether two variables are statistically independent.

If two variables take the same categorical values, the table is called a "square" contingency table. When the observed frequencies are concentrated in the main diagonal cells, the two variables are dependent. Even if the observations are not concentrated on the main diagonal but we have one large frequency and several small frequencies in each row and each column, then there is a strong association between the categories of a variable and those of the other, and hence a strong dependence. This is a common situation in real world data and, since the case of independence is infrequent and unrealistic, a suitable model for representing dependence data is important . Consequently, many statisticians consider various statistical models instead of an independence model and study the method of estimation and hypothesis testing based on a statistical model.When statistical independence between two variables does not hold, association models, which indicate the structure of odds ratios, have been considered to analyze contingency tables. On the other hand, symmetry or asymmetry models, which indicate the structure of ratios for cell probabilities in symmetric positions, are often used to analyze square contingency tables.

This study proposes a model with characteristics of both an association model and asymmetry model. Our model is more parsimonious than many association or asymmetry models. Hence, our model may better estimate the distribution than conventional association models and asymmetry models.

This paper is organized as follows. Section 2 introduces previous research and proposes an asymmetry plus association model. Section 3 describes the necessary and sufficient condition to use our model. Section 4 provides the methods to evaluate model-fitting based on goodness-of-fit. Section 5 concludes this paper.

2. Models

For an $r \times r$ square contingency table with ordinal categories, let π_{ij} denote the probability that an observation will fall in the ith row and jth column of the contingency table ($i = 1, \ldots, r; j = 1, \ldots, r$). Goodman [3–5] considered many association models in a contingency table. For example, the quasi-uniform association (QU) model is defined as

$$\pi_{ij} = \begin{cases} \mu\alpha_i\beta_j\theta^{ij} & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases} \tag{1}$$

Without loss of generality, we impose $\alpha_r = \beta_r = 1$. The odds ratio for rows i and j (>i), and columns s and t (>s) are denoted by $\phi_{(ij;st)}$. That is,

$$\phi_{(ij;st)} = \frac{\pi_{is}\pi_{jt}}{\pi_{js}\pi_{it}}. \tag{2}$$

Using the odds ratios, the QU model can be expressed as

$$\phi_{(ij;st)} = \theta^{(j-i)(t-s)} \quad (i \neq s, i \neq t, j \neq s, j \neq t). \tag{3}$$

The QU model with $\theta = 1$ is the quasi-independence (QI) model (see p. 426 in Agresti [6]). That is,

$$\pi_{ij} = \begin{cases} \mu\alpha_i\beta_j & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases} \tag{4}$$

On the other hand, many statisticians have analyzed square contingency tables using a symmetric structure or an asymmetric structure for cell probabilities. Bowker [7] proposed the symmetry (S) model, which is defined as

$$\pi_{ij} = \psi_{ij} \quad (i = 1, \ldots, r; j = 1, \ldots, r), \tag{5}$$

where $\psi_{ij} = \psi_{ji}$. This model indicates the symmetric structure for cell probabilities.

Stuart [8] proposed the marginal homogeneity (MH) model, which is defined as

$$\pi_{i+} = \pi_{+i} \quad (i = 1, \ldots, r), \tag{6}$$

where $\pi_{i+} = \sum_{j=1}^{r} \pi_{ij}$ and $\pi_{+i} = \sum_{j=1}^{r} \pi_{ji}$. The MH model indicates that the row marginal distribution is equivalent to the column marginal distribution.

Caussinus [9] proposed the quasi-symmetry (QS) model, which is defined as

$$\pi_{ij} = \mu\alpha_i\beta_j\psi_{ij} \quad (i = 1, \ldots, r; j = 1, \ldots, r), \tag{7}$$

where $\psi_{ij} = \psi_{ji}$. This model is identical to the S model when $\alpha_i = \beta_i$. The QS model can be expressed as

$$\phi_{(ij;st)} = \phi_{(st;ij)} \quad (i < j; s < t). \tag{8}$$

The QS model indicates the symmetric structure of the odds ratios. The QU model implies the QS model. That is, the QS model holds when the QU model holds.

When the S model does not hold, asymmetry models, with a weaker restriction than the S model, have been proposed. For example, Tahata and Tomizawa [10] proposed the kth linear asymmetry (LS_k) model, which is defined for a fixed k $(k = 1, \ldots, r-1)$ as

$$\pi_{ij} = \mu \prod_{l=1}^{k} \alpha_l^{i^l} \beta_l^{j^l} \psi_{ij} \quad (i = 1, \ldots, r; j = 1, \ldots, r), \tag{9}$$

where $\psi_{ij} = \psi_{ji}$. Note that when $\alpha_l = \beta_l$, this model is the S model. As k increases, the LS_k model is less restrictive, and the LS_{r-1} model is the QS model. Namely, the LS_k model is the intermediate model between the S model and QS model. The LS_k model can be expressed as

$$\frac{\pi_{ij}}{\pi_{ji}} = \prod_{l=1}^{k} \gamma_l^{j^l - i^l} \quad (i \neq j). \tag{10}$$

The LS_k model includes the linear diagonals-parameter symmetry model [11] and the extended linear diagonals-parameter symmetry model [12].

Goodman [4] introduced the symmetry plus quasi-independence (SQI) model, which is defined as

$$\pi_{ij} = \begin{cases} \mu \alpha_i \alpha_j & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases} \tag{11}$$

This model is a special case of the S model and the QI model when $\psi_{ij} = \mu \alpha_i \alpha_j$ and $\alpha_i = \beta_i$ for $i \neq j$, respectively.

Yamamoto and Tomizawa [13] proposed the symmetry plus quasi-uniform association (SQU) model, which is defined as

$$\pi_{ij} = \begin{cases} \mu \alpha_i \alpha_j \theta^{ij} & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases} \tag{12}$$

The SQU model implies the S model and QU model. Note that the SQU model is identical to the SQI model when $\theta = 1$.

Association models and asymmetry models have been proposed independently. However, an asymmetry plus association model, which considers both the structure of asymmetry for cell probabilities and odds ratios, is rarely considered.

Here, we propose a new model defined for a fixed k $(k = 1, \ldots, r-1)$ as

$$\pi_{ij} = \begin{cases} \mu \alpha_i \alpha_j \prod_{l=1}^{k} \delta_l^{j^l - i^l} \theta^{ij} & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases} \tag{13}$$

Without loss of generality, we set $\alpha_r = 1$. This model is called the kth linear asymmetry plus quasi-uniform association (LSQU_k) model. When $\theta = 1$, it is called the kth linear asymmetry plus quasi-independence (LSQI_k) model.

If the LSQU_k model holds, then

$$\frac{\pi_{ij}}{\pi_{ji}} = \prod_{l=1}^{k} \delta_l^{2(j^l - i^l)} \quad (i \neq j). \tag{14}$$

The LS_k model holds by $\gamma_l = \delta_l^2$ in Equation (14). Additionally,

$$\phi_{(ij;st)} = \theta^{(j-i)(t-s)} \quad (i \neq s, i \neq t, j \neq s, j \neq t). \tag{15}$$

Therefore, the LSQU_k model shows characteristics of both the LS_k model and the QU model.

This model with $\delta_l = 1$ for $l = 1, \ldots, k$ is the SQU model. When $k = r - 1$, the LSQU$_k$ model implies

$$\frac{\pi_{ij}}{\pi_{ji}} = \prod_{l=1}^{r-1} \frac{\gamma_l^{j^l}}{\gamma_l^{i^l}} \quad (i \neq j). \tag{16}$$

On the other hand, the QU model implies

$$\frac{\pi_{ij}}{\pi_{ji}} = \frac{\lambda_j}{\lambda_i} \quad (i \neq j), \tag{17}$$

where $\lambda_j = \beta_j / \alpha_j$. Setting $\lambda_j = \prod_{l=1}^{r-1} \gamma_l^{j^l}$ provides a one-to-one relation between $\{\lambda_1, \ldots, \lambda_{r-1}\}$ and $\{\gamma_1, \ldots, \gamma_{r-1}\}$. This means that the LSQU$_{r-1}$ model is equivalent to the QU model. The LSQU$_k$ ($k < r - 1$) model is a special case of the QU model since the LSQU$_{r-1}$ model with $\delta_l = 1$ for $l = k + 1, \ldots, r - 1$ is the LSQU$_k$ model. Hence, the LSQU$_k$ model is an intermediate model between the SQU and QU models. Similarly, the LSQI$_{r-1}$ model is equivalent to the QI model. That is, the LSQI$_k$ model is an intermediate model between the SQI and QI models (For more details, see Figure 1).

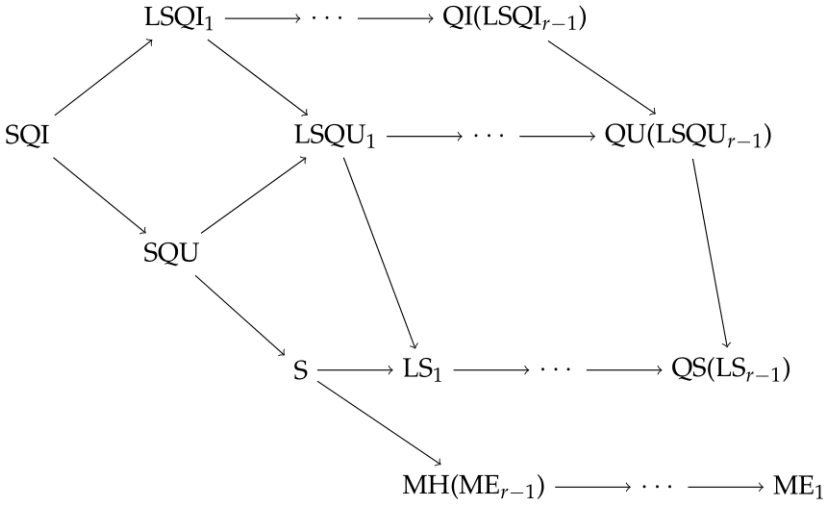

Figure 1. Relationships among the models (A → B indicates that model A is a special case of model B).

3. Necessary and Sufficient Condition for the SQU Model

Caussinus [9] introduced the necessary and sufficient condition for the S model. This condition separates the S model into multiple models with a weaker restriction than the S model. Assuming that model M_1 holds if and only if both models M_2 and M_3 hold, then analyzing models M_2 and M_3 should elucidate a more detailed structure of the cell probabilities. Here, we are interested in deriving a necessary and sufficient condition for the SQU model using the LSQU$_k$ model.

Yamamoto and Tomizawa [13] provided the following necessary and sufficient condition for the SQU model.

Theorem 1. *The SQU model holds if and only if both the QU model and the MH model hold.*

Let X and Y denote the row and column variables, respectively, and consider a model defined for a fixed k ($k = 1, \ldots, r - 1$), which is given as

$$E(X^l) = E(Y^l) \quad (l = 1, \ldots, k), \tag{18}$$

where $E(X^l) = \sum_i \sum_j i^l \pi_{ij}$ and $E(Y^l) = \sum_i \sum_j j^l \pi_{ij}$. This model can be referred to as the marginal kth moment equality (ME$_k$) model. This leads to the following theorem.

Theorem 2. *For any k $(k = 1, \ldots, r-1)$, the SQU model holds if and only if both the LSQU$_k$ model and the ME$_k$ model hold.*

Proof. If the SQU model holds, the LSQU$_k$ model holds because the LSQU$_k$ model with $\delta_l = 1$ $(l = 1, \ldots, k)$ is the SQU model. Since the SQU model implies the S model, we can see that

$$E(X^l) = \sum_i \sum_j i^l \pi_{ij} = \sum_i \sum_j i^l \pi_{ji} = E(Y^l) \quad (l = 1, \ldots, k). \tag{19}$$

The ME$_k$ model also holds. The necessity is proved.

Conversely, if both the LSQU$_k$ model and the ME$_k$ model hold, we can prove that the SQU model holds. If the LSQU$_k$ model holds, from Equation (14), we obtain

$$\log \pi_{ij} - \log \pi_{ji} = 2 \sum_{l=1}^{k} (j^l - i^l) \log \delta_l \quad (i \neq j). \tag{20}$$

The ME$_k$ model can also be expressed as

$$\sum_{i \neq j} \sum (j^l - i^l) \pi_{ij} = 0 \quad (l = 1, \ldots, k). \tag{21}$$

From the LSQU$_k$ model and the ME$_k$ model, we obtain

$$\begin{aligned} \sum_{i \neq j} \sum (\pi_{ij} - \pi_{ji})(\log \pi_{ij} - \log \pi_{ji}) &= 2 \sum_{l=1} \log \delta_l \sum_{i \neq j} \sum (j^l - i^l)(\pi_{ij} - \pi_{ji}) \\ &= 0. \end{aligned} \tag{22}$$

Since the logarithmic function is strictly increasing, then for any $i \neq j$

$$(\pi_{ij} - \pi_{ji})(\log \pi_{ij} - \log \pi_{ji}) \geq 0. \tag{23}$$

Equation (22) with $\pi_{ij} = \pi_{ji}$ holds, that is, the S model holds. When the S model holds, the MH model holds. Additionally, the LSQU$_k$ model is a special case of the QU model. From Theorem 1, the SQU model holds. The proof is complete. □

Theorem 2 is a generalization of Yamamoto and Tomizawa's result because the ME$_{r-1}$ model is equivalent to the MH model (see [14]). This leads to the following corollary.

Corollary 1. *For any k $(k = 1, \ldots, r-1)$, the SQI model holds if and only if both the LSQI$_k$ model and the ME$_k$ model hold.*

4. Partition of Test Statistics

Here, we describe a method to evaluate the model fitting. We consider a test of hypothesis, where the null hypothesis is that model M holds, and the alternative hypothesis is that model M does not hold. Let n_{ij} denote the observed frequency in the (i, j)th cell of the table and m_{ij} indicate the corresponding expected frequency with $n = \sum_i \sum_j n_{ij}$ $(i = 1, \ldots, r; j = 1, \ldots, r)$. Assume that $\{n_{ij}\}$ has a multinomial distribution. Then $\hat{m}_{ij}$ denotes the maximum likelihood estimate (MLE) of m_{ij} under a model. The likelihood ratio chi-squared statistic for the goodness-of-fit of the model M is defined as

$$G^2(M) = 2 \sum_{i=1}^{r} \sum_{j=1}^{r} n_{ij} \log \left(\frac{n_{ij}}{\hat{m}_{ij}} \right). \tag{24}$$

The numbers of degrees of freedom (df) for testing the goodness-of-fit under the SQU, LSQU$_k$, and ME$_k$ models are $r^2 - 2r - 1$, $r^2 - 2r - 1 - k$, and k, respectively. The number of df for the SQU model is equal to the sum of those for the LSQU$_k$ and ME$_k$ models.

Previous studies have discussed the separability of a model [15–19]. Separability means that a test statistic for the goodness-of-fit of model M_1 is asymptotically equivalent to the sum of the test statistics for model M_2 and model M_3 when model M_1 can be separated into model M_2 and model M_3. If it holds, the incompatible situation, where both model M_2 and model M_3 are accepted but model M_1 is rejected, would not arise. This leads to the following theorem.

Theorem 3. *For any k ($k = 1, \ldots, r-1$), the test statistic $G^2(SQU)$ is asymptotically equivalent to the sum of $G^2(LSQU_k)$ and $G^2(ME_k)$.*

Proof. For a fixed k ($k = 1, \ldots, r-1$), the LSQU_k model can be expressed as

$$\log \pi_{ij} = \begin{cases} \mu' + \alpha'_i + \alpha'_j + \sum_{l=1}^{k} (j^l - i^l)\delta'_l + ij\theta' & (i \neq j), \\ \mu' + \alpha'_i + \alpha'_i + \psi'_{ii} & (i = j). \end{cases} \tag{25}$$

Without loss of generality, we can impose $\alpha'_r = 0$. Let

$$\pi = (\pi_{11}, \ldots, \pi_{1r}, \pi_{21}, \ldots, \pi_{2r}, \ldots, \pi_{rr})^T, \tag{26}$$

and

$$\beta = (\mu', \beta_1, \beta_2, \beta_{12})^T, \tag{27}$$

where "T" denotes the transpose,

$$\beta_1 = (\alpha'_1, \ldots, \alpha'_{r-1}), \quad \beta_2 = (\delta'_1, \ldots, \delta'_k), \tag{28}$$

and

$$\beta_{12} = (\theta', \psi'_{11}, \ldots, \psi'_{rr}). \tag{29}$$

The LSQU_k model can also be expressed as

$$\log \pi = X\beta = (1_{r^2}, X_1, X_2, X_{12})\beta, \tag{30}$$

where $\log \pi = (\log \pi_{11}, \ldots, \log \pi_{rr})^T$, X is the $r^2 \times (2r + 1 + k)$ matrix, and 1_s is the $s \times 1$ vector of the 1 element. Additionally,

$$X_1 = \begin{pmatrix} I_{r-1} \otimes 1_r \\ O_{r,r-1} \end{pmatrix} + 1_r \otimes \begin{pmatrix} I_{r-1} \\ 0_{r-1}^T \end{pmatrix} \tag{31}$$

$$X_2 = (x_1, \ldots, x_k), \tag{32}$$

where

$$x_l = 1_r \otimes J_r^l - J_r^l \otimes 1_r \quad (l = 1, \ldots, k), \tag{33}$$

and X_{12} is the $r^2 \times (r+1)$ matrix determined from Equation (25). Note that O_{st} is the $s \times t$ zero matrix, 0_s is the $s \times 1$ zero vector, $J_r^l = (1^l, \ldots, r^l)^T$, and "$\otimes$" represents the Kronecker product. The matrix X has a full column rank, which is $K = 2r + 1 + k$.

We denote the linear space spanned by the columns of the matrix X by $S(X)$ with dimension K. Let U be an $r^2 \times d_1$ full column rank matrix, where $d_1 = r^2 - 2r - 1 - k$, such that $S(U)$ is the orthogonal complement of space $S(X)$. Hence, $U^T X = O_{d_1,K}$.

Let $h_1(\pi)$ be a vector of functions defined by $h_1(\pi) = U^T \log \pi$. Moreover, let $h_2(\pi)$ be a vector of functions defined by $h_2(\pi) = X_2^T \pi$, and note that $X_2^T U = O_{d_2,d_1}$ where $d_2 = k$ because X_2 belongs to space $S(X)$.

From Equation (25), the LSQU_k model is equivalent to the hypothesis $h_1(\pi) = 0_{d_1}$. Additionally, the ME_k model is equivalent to the hypothesis $h_2(\pi) = 0_{d_2}$. From Theorem 2, the SQU model is equivalent to the hypothesis $h_3(\pi) = 0_{d_3}$ where $h_3 = (h_1^T, h_2^T)^T$ and $d_3 = d_1 + d_2 = r^2 - 2r - 1$.

We derive the Wald statistic for the SQU model in an analogous mannar to Bhapkar [20]. Let H_s $(s = 1, 2, 3)$ denote the $d_s \times r^2$ matrix of partial derivatives of $h_s(\pi)$ with respect to π. Namely, $H_s(\pi) = \partial h_s(\pi)/\partial \pi^T$. Let $\Sigma(\pi) = diag(\pi) - \pi\pi^T$, where $diag(\pi)$ denotes a diagonal matrix with the ith component of π as the ith diagonal component. Additionally, let p_{ij} denote a sample proportion of the (i, j) cell. That is, $p_{ij} = n_{ij}/n$, and $p = (p_{11}, \ldots, p_{1r}, p_{21}, \ldots, p_{2r}, \ldots, p_{rr})^T$. The central limit theorem indicates that $\sqrt{n}(p - \pi)$ has an asymptotic normal distribution with mean 0_{r^2} and covariance matrix $\Sigma(\pi)$. Using the delta method, $\sqrt{n}(h_3(p) - h_3(\pi))$ has an asymptotic normal distribution with mean 0_{d_3} and covariance matrix

$$H_3(\pi)\Sigma(\pi)H_3^T(\pi) = \begin{pmatrix} H_1(\pi)\Sigma(\pi)H_1^T(\pi) & H_1(\pi)\Sigma(\pi)H_2^T(\pi) \\ H_2(\pi)\Sigma(\pi)H_1^T(\pi) & H_2(\pi)\Sigma(\pi)H_2^T(\pi) \end{pmatrix}.$$

Since $H_1(\pi)\pi = U^T 1_{r^2} = 0_{d_1}$, $H_1(\pi)diag(\pi) = U^T$, and $H_2(\pi) = X_2^T$, we obtain

$$H_1(\pi)\Sigma(\pi)H_2^T(\pi) = U^T X_2 = O_{d_1,d_2}. \tag{34}$$

Under each hypothesis, $h_s(\pi) = 0_{d_s}$ $(s = 1, 2, 3)$, we see

$$W_3 = W_1 + W_2, \tag{35}$$

where

$$W_s = nh_s(p)^T(H_s(p)\Sigma(p)H_s^T(p))^{-1}h_s(p). \tag{36}$$

The Wald statistic W_s has an asymptotic chi-squared distribution with d_s df. That is, (i) W_1 is the Wald statistic for the LSQU_k model, (ii) W_2 is that for the ME_k model, and (iii) W_3 is that for the SQU model. The proof is completed using the asymptotic equivalence of the Wald statistic and the likelihood ratio statistic as proved by Rao [21]. □

Theorem 3 is also a generalization of Yamamoto and Tomizawa's result since this theorem is identical to Yamamoto and Tomizawa's result when $k = r - 1$. Moreover, we obtain the following corollary.

Corollary 2. *For any k $(k = 1, \ldots, r - 1)$, the test statistic $G^2(SQI)$ is asymptotically equivalent to the sum of $G^2(LSQI_k)$ and $G^2(ME_k)$.*

5. Example

Table 1 shows the data cited by [22]. This data described 59 matched pairs using 4 dose levels of conjugated estrogen. The models described herein are used to analyze this data. Table 2 shows the value of $G^2(M)$ for each model applied to the data in Table 1. That is, for model M, the null hypothesis is that model M holds, and the alternative hypothesis is that model M does not hold. From Table 2, the SQI, SQU, S, and ME_k models do not fit well, and the LSQI_k, LSQU_k, and LS_k models are accepted at the 0.05 significant level $(k = 1, 2, 3)$. We choose the most appropriate model in these models. If model M_1 is a special case of model M_2, a test based on the difference between the likelihood ratio chi-squared statistic can compare the model fitting of two nested models. Let d_1 and d_2 denote the degrees of freedom for the models M_1 and M_2, respectively. Assuming that model M_2 holds, a likelihood ratio chi-squared statistic under model M_1 is given as $G^2(M_1 \mid M_2) = G^2(M_1) - G^2(M_2)$. This statistic is an asymptotically chi-squared distribution with $d_1 - d_2$ degrees of freedom. When we use it at the 0.05 significant level, the LSQI_1 model is the most appropriate model.

Table 1. Average doses of conjugated estrogen used by cases and matched controls: Los Angeles endometrial cancer study [22].

Average Dose for Case (mg/day)	Average Dose for Control (mg/day)				
	0 (1)	**0.1–0.299 (2)**	**0.3–0.625 (3)**	**0.625+ (4)**	**Total**
0 (1)	6	2	3	1	12
0.1–0.299 (2)	9	4	2	1	16
0.3–0.625 (3)	9	2	3	1	15
0.625+ (4)	12	1	2	1	16
Total	36	9	10	4	59

Table 2. The values of likelihood ratio chi-squared statistics for models applied to Table 1.

Model	**df**	$G^2(M)$
SQI	8	19.98 *
SQU	7	19.86 *
$LSQI_1$	7	3.62
$LSQI_2$	6	2.98
$LSQI_3$ (QI)	5	0.77
$LSQU_1$	6	3.61
$LSQU_2$	5	2.98
$LSQU_3$ (QU)	4	0.69
S	6	19.27 *
LS_1	5	2.97
LS_2	4	2.33
LS_3 (QS)	3	0.46
ME_1	1	16.43 *
ME_2	2	17.08 *
ME_3 (MH)	3	19.12 *

Note * Significant at the 0.05 level.

Table 3 shows the estimated expected frequencies from the $LSQI_1$ model for the data in Table 1. The value of maximum likelihood estimator of δ_1 for the $LSQI_1$ model is 0.71. We estimate the ratio between two probabilities as $\hat{\pi}_{ij}/\hat{\pi}_{ji} = 0.71^{2(j-i)}$ for $i < j$. Therefore, the probability distribution for the average dose for a case tends to be stochastically higher than the probability distribution for the average dose for control because $\hat{\delta}_1 < 1$.

Table 3. Estimated expected frequencies from the $LSQI_1$ model.

Average Dose for Case (mg/day)	Average Dose for Control (mg/day)			
	0 (1)	**0.1–0.299 (2)**	**0.3–0.625 (3)**	**0.625+ (4)**
0 (1)	6	3.58	2.64	1.42
0.1–0.299 (2)	7.07	4	1.13	0.61
0.3–0.625 (3)	10.34	2.24	3	0.89
0.625+ (4)	10.96	2.37	1.76	1

Finally, we are interested in inferring the reason for the poor fit of the SQI model. According to Corollary 1, the SQI model is separated into the $LSQI_1$ model and the ME_1 model. Since the $LSQI_1$ model fits very well, but the ME_1 model fits very poorly, we deduce that the lack of structure of the ME_1 model is responsible for the poor fit of the SQI model.

6. Conclusions

Herein we describe an asymmetry plus association model. This model indicates the asymmetry structures for cell probabilities between symmetric position and odds ratios. Our model is an intermediate model between the SQU model and the QU model. If the

QU ($LSQU_{r-1}$) model holds but the SQU model does not, the $LSQU_k$ model for $k < r-1$ may hold. In this case, the QU model may be overfitting. That is, our model may realize a better fit than the QU model under these conditions. In practice, the $LSQI_1$ model fits well when the SQU model fits poorly and the QU model fits for the data in Table 1. Additionally, a theorem with respect to the necessary and sufficient condition for the SQU model is represented using our model. Using this theorem, we show the asymptotic separability for the SQU model. Namely, the likelihood ratio chi-squared statistic for the SQU model is equivalent to the sum of those for the separated models, which helps deduce the reason that the SQU model does not hold.

Author Contributions: Conceptualization, K.T.; methodology, K.F.; formal analysis, K.F; writing—original draft preparation, K.F.; writing—review and editing, K.F. and K.T.; project administration, K.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank the three anonymous referees for their helpful comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

QU	Quasi-uniform association
QI	Quasi-independence
S	Symmetry
MH	Marginal homogeneity
QS	Quasi-symmetry
LS_k	kth linear asymmetry
SQI	Symmetry plus quasi-independence
SQU	Symmetry plus quasi-uniform association
$LSQU_k$	kth linear asymmetry plus quasi-uniform association
$LSQI_k$	kth linear asymmetry plus quasi-independence
ME_k	Marginal kth moment equality
df	Degrees of freedom

References

1. Bishop, Y.M.M.; Fienberg, S.E.; Holland, P.W. *Discrete Multivariate Analysis: Theory and Practice*; The MIT Press: Cambridge, MA, USA, 1975.
2. Kateri, M. *Contingency Table Analysis: Methods and Implementation Using R*; Birkhäuser(Springer): New York, NY, USA, 2014.
3. Goodman, L.A. Simple models for the analysis of association in cross-classifications having ordered categories. *J. Am. Stat. Assoc.* **1979**, *74*, 537–552. [CrossRef]
4. Goodman, L.A. The analysis of cross-classified data having ordered and/or unordered categories: Association models, correlation models, and asymmetry models for contingency tables with or without missing entries. *Ann. Stat.* **1985**, *13*, 10–69. [CrossRef]
5. Goodman, L.A. Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables. *Int. Stat. Rev.* **1986**, *54*, 243–309. [CrossRef]
6. Agresti, A. *Categorical Data Analysis*, 2nd ed.; Wiley: New York, NY, USA, 2002.
7. Bowker, A.H. A test for symmetry in contingency tables. *J. Am. Stat. Assoc.* **1948**, *43*, 572–574. [CrossRef] [PubMed]
8. Stuart, A. A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika* **1955**, *42*, 412–416. [CrossRef]
9. Caussinus, H. Contribution à l'analyse statistique des tableaux de corrélation. *Ann. Fac. Des. Sci. L'Univ. Toulouse* **1965**, *29*, 77–182. [CrossRef]
10. Tahata, K.; Tomizawa, S. Generalized linear asymmetry model and decomposition of symmetry for multiway contingency tables. *J. Biom. Biostat.* **2011**, *2*, 1–6. [CrossRef]
11. Agresti, A. A simple diagonals-parameter symmetry and quasi-symmetry model. *Stat. Probab. Lett.* **1983**, *1*, 313–316. [CrossRef]

12. Tomizawa, S. An extended linear diagonals-parameter symmetry model for square contingency tables with ordinal categories. *Metron* **1991**, *49*, 401–409.
13. Yamamoto, K.; Tomizawa, S. Symmetry plus quasi uniform association model and its orthogonal decomposition for square contingency tables. *J. Mod. Appl. Stat. Methods* **2010**, *9*, 255–262. [CrossRef]
14. Tahata, K.; Tomizawa, S. Generalized marginal homogeneity model and its relation to marginal equimoments for square contingency tables with ordered categories. *Adv. Data Anal. Classif.* **2008**, *2*, 295–311. [CrossRef]
15. Aitchison, J. Large-sample restricted parametric tests. *J. R. Stat. Soc. Ser. B* **1962**, *24*, 234–250. [CrossRef]
16. Darroch, J.N.; Silvey, S.D. On testing more than one hypothesis. *Ann. Math. Stat.* **1963**, *34*, 555–567. [CrossRef]
17. Lang, J.B.; Agresti, A. Simultaneously modeling joint and marginal distributions of multivariate categorical responses. *J. Am. Stat. Assoc.* **1994**, *89*, 625–632. [CrossRef]
18. Lang, J.B. On the partitioning of goodness-of-fit statistics for multivariate categorical response models. *J. Am. Stat. Assoc.* **1996**, *91*, 1017–1023. [CrossRef]
19. Tomizawa, S.; Tahata, K. The analysis of symmetry and asymmetry: Orthogonality of decomposition of symmetry into quasi-symmetry and marginal symmetry for multi-way tables. *J. Soc. Fr. Stat.* **2007**, *148*, 3–36.
20. Bhapkar, V.P. A note on the equivalence of two test criteria for hypotheses in categorical data. *J. Am. Stat. Assoc.* **1966**, *61*, 228–235. [CrossRef]
21. Rao, C.R. *Linear Statistical Inference and Its Applications*, 2nd ed.; Wiley: New York, NY, USA, 1973.
22. Breslow, N.E.; Day, N.E. *Statistical Methods in Cancer Research, Vol. I: The Analysis of Case-Control Studies*; International Agency for Research on Cancer: Lyon, France, 1980.