

Decompositions for Ordinal Quasi-Symmetry Model in Square Contingency Tables with Ordered Categories

¹Kouji Yamamoto, ²Satoru Shinoda and ³Sadao Tomizawa
¹Center for Clinical Investigation and Research, Osaka University Hospital,
 2-15, Yamadaoka, Suita, Osaka, 565-0871, Japan
^{2,3}Department of Information Sciences, Faculty of Science and Technology,
 Tokyo University of Science, Noda City, Chiba, 278-8510, Japan

Abstract: Problem statement: For square contingency tables with ordered categories, this study considers four kinds of extensions of the marginal homogeneity model and gives decompositions for the ordinal quasi-symmetry model. The decompositions are extensions of some existing decompositions. **Approach:** This study gives a decomposition theorem that the ordinal quasi-symmetry model holds if and only if the quasi-symmetry model and the proposed weighted marginal homogeneity model hold. An example is given. **Results:** For the data of cross-classification of father's and his son's occupational status in Denmark, the decomposition of the ordinal quasi-symmetry model is applied and the detailed analysis is given. **Conclusion:** When the ordinal quasi-symmetry model fits the data poorly, this decomposition is useful for seeing which of decomposed two models influences stronger.

Key words: Linear diagonals-parameter symmetry, score, quasi-symmetry, weighted marginal homogeneity

INTRODUCTION

Consider an $R \times R$ square contingency tables with the same row and column classifications. Let p_{ij} denote the probability that an observation will fall in the i th row and j th column of the table ($i = 1, \dots, R; j = 1, \dots, R$). The Quasi-Symmetry (QS) model (Causinus, 1965) is defined by:

$$p_{ij} = \mu \alpha_i \beta_j \psi_{ij} \quad (i = 1, \dots, R; j = 1, \dots, R)$$

where, $\psi_{ij} = \psi_{ji}$. A special case of this model with $\{\alpha_i = \beta_i\}$ is the usual Symmetry (S) model (Bowker, 1948).

The QS model can be also expressed in a log-linear form as (Bishop *et al.*, 1978):

$$\log p_{ij} = \nu + \nu_{1(i)} + \nu_{2(j)} + \nu_{12(ij)} \\ (i = 1, \dots, R; j = 1, \dots, R)$$

Where:

$$\nu_{12(ij)} = \nu_{12(ji)} \\ \sum_{k=1}^R \nu_{1(k)} = \sum_{k=1}^R \nu_{2(k)} = \sum_{k=1}^R \nu_{12(kj)} = \sum_{k=1}^R \nu_{12(ik)} = 0$$

The Marginal Homogeneity (MH) model is defined by:

$$p_{i\cdot} = p_{\cdot i} \quad (i = 1, \dots, R)$$

where, $p_{i\cdot} = \sum_{t=1}^R p_{it}$ and $p_{\cdot i} = \sum_{s=1}^R p_{si}$ (Stuart, 1955).

Causinus (1965) gave the decomposition of the S model as follows:

Theorem 1: The S model holds if and only if both the QS and MH models hold.

Each of the S, QS and MH models is invariant under arbitrary permutations of same row and column categories. So, these models are suitable for analyzing the data on a nominal scale, but they are possible for analyzing the data on an ordinal scale because they only require a categorical scale.

The Ordinal Quasi-Symmetry (OQS) model (Agresti, 2002), which is a special case of the QS model, is often useful when the categories are ordered. Let $u_1 < \dots < u_R$ denote the ordered known scores assigned for both the rows and columns of same classifications. The OQS model is defined by:

Corresponding Author: Kouji Yamamoto, Center for Clinical Investigation and Research, Osaka University Hospital, 2-15, Yamadaoka, Suita, Osaka, 565-0871, Japan

$$p_{ij} = \begin{cases} \delta^{u_i - u_j} \psi_{ij} & (i < j) \\ \psi_{ij} & (i \geq j) \end{cases}$$

where, $\psi_{ij} = \psi_{ji}$. The OQS model with scores $\{i\}$ instead of $\{u_i\}$ is the Linear Diagonals-Parameter Symmetry (LDPS) model (Agresti, 1983). Thus, the LDPS model is useful for analyzing the data with positive integer scores or equal-interval scores. While the OQS model is useful for analyzing the data with any increasing (or decreasing) scores.

Tomizawa (1987) introduced two kinds of Diagonal Weighted Marginal homogeneity (DWM) models and gave the decompositions of the LDPS model as follows.

Theorem 2: The LDPS model holds if and only if both the QS and DWM models hold.

Consider the data in Table 1, taken directly from Bishop *et al.* (1978), which represents father's and his son's occupational status in Denmark. Each category means that (1) is High professionals, (2) White-collar employees of higher education, (3) White-collar employees of less high education, (4) Upper working class and (5) Unskilled workers. We treat the classifications as ordinal.

First, we apply the LDPS model to these data. For testing goodness-of-fit of the LDPS model, the value of likelihood ratio statistic (denoted by G^2) is 19.054 with 9 degrees of freedom ($p = 0.025$). Thus, the LDPS model does not fit these data well. This may be because we assign the equal-interval scores for the categories.

Next we apply the OQS model to the data in Table 1. We shall assign the same ordered scores for both the rows and columns of same classifications by using empirical ridit. Then, we use the average of ridit for a row category and that for the corresponding column category. Therefore, we assign the scores $u_1 = 0.014$, $u_2 = 0.089$, $u_3 = 0.293$, $u_4 = 0.604$ and $u_5 = 0.886$ for both the rows and columns of same classifications. Then the G^2 value for the OQS model is 16.654 with 9 degrees of freedom ($p = 0.054$). Thus, the OQS model fits these data moderately.

Table 1: Occupational status for Danish father-son pairs; from Bishop *et al.* (1978)

Father's status	Son's status					Total
	(1)	(2)	(3)	(4)	(5)	
(1)	18	17	16	4	2	57
(2)	24	105	109	59	21	318
(3)	23	84	289	217	95	708
(4)	8	49	175	348	198	778
(5)	6	8	69	201	246	530
Total	79	263	658	829	562	2391

Also, we apply the QS model to these data. Since the value of G^2 is 6.468 with 6 degrees of freedom ($p = 0.373$), the QS model also fits these data well. Note that each of the OQS and LDPS models is a special case of the QS model.

Then, since the OQS model implies the QS model and these models fit the data in Table 1 well, we consider which of the models fits these data much better. For testing the goodness-of-fit of the OQS model under the assumption that the QS model holds true, the difference between the G^2 values for the OQS and QS models is 10.186 with 3 degrees of freedom ($p = 0.017$). Therefore, we see that the OQS model does not fit these data well, assuming that the QS model holds true.

Therefore, we shall consider a decomposition of the OQS model, to investigate the reason that the OQS model has a poor fit for the data in Table 1.

The purpose of this study is to give a decomposition of the OQS model, introducing four kinds of new models.

MATERIALS AND METHODS

To consider decompositions of the OQS model, we shall introduce four kinds of new models. First, consider a model defined by:

$$p_i^-(\delta_1) + p_{ii} + p_i^+ = p_i^+ + p_{ii} + p_i^-(\delta_1) \quad (i = 1, \dots, R)$$

Where:

$$p_i^+ = \sum_{k=i+1}^R p_{ik}, \quad p_i^- = \sum_{k=1}^{i-1} p_{ki}$$

$$p_i^-(\delta_1) = \sum_{k=1}^{i-1} \delta_1^{u_i - u_k} p_{ik}, \quad p_i^+(\delta_1) = \sum_{k=i+1}^R \delta_1^{u_k - u_i} p_{ki}$$

This model indicates that the row marginal totals summed by multiplying the probabilities p_{ij} for the cells with a distance $i-j (>0)$ below main diagonal in the table by the weight $\delta_1^{u_i - u_j} (>0)$ are equal to the column marginal totals summed by the same way. We will refer to this model as the weighed marginal homogeneity 1 (WMH-1) model.

Next, consider the following model defined by:

$$p_i^-(\delta_2) + p_{ii} + p_i^+(\delta_2) = p_i^+(\delta_2) + p_{ii} + p_i^-(\delta_2) \quad (i = 1, \dots, R)$$

Where:

$$p_i^- = \sum_{k=1}^{i-1} p_{ik}, \quad p_i^+ = \sum_{k=i+1}^R p_{ki}$$

$$p_i^-(\delta_2) = \sum_{k=i+1}^R \delta_2^{u_k - u_i} p_{ik}, \quad p_i^+(\delta_2) = \sum_{k=1}^{i-1} \delta_2^{u_i - u_k} p_{ki}$$

This model states that the row marginal totals summed by multiplying the probabilities p_{ij} for the cells with a distance $j-i (> 0)$ above main diagonal in the table by the weight $\delta_2^{u_i - u_j} (> 0)$ are equal to the column marginal totals summed by the same way. This is contrast to the WMH-1 model. We will refer to this model as the weighed marginal homogeneity 2 (WMH-2) model.

Thirdly, we introduce a model defined by:

$$p_i^*(\delta_3) = p_i^*(\delta_3) \quad (i = 1, \dots, R)$$

Where:

$$p_i^*(\delta_3) = \sum_{k=1}^R \delta_3^{u_k} p_{ik}, \quad p_i^*(\delta_3) = \delta_3^{u_i} \sum_{k=1}^R p_{ki}$$

This model shows that the i th row marginal totals summed by multiplying the probabilities for the cells in the k th column by the weight $\delta_3^{u_k} (> 0)$ is equal to the i th column marginal totals multiplied by $\delta_3^{u_i}$. We will refer to this model as the weighed marginal homogeneity 3 (WMH-3) models.

Finally, consider a model:

$$p_i^{**}(\delta_4) = p_i^{**}(\delta_4) \quad (i = 1, \dots, R)$$

Where:

$$p_i^{**}(\delta_4) = \delta_4^{u_i} \sum_{k=1}^R p_{ik}, \quad p_i^{**}(\delta_4) = \sum_{k=1}^R \delta_4^{u_k} p_{ki}$$

This model represents that the i th row marginal totals multiplied by $\delta_4^{u_i} (> 0)$ is equal to the i th column marginal totals summed by multiplying the probabilities for the cells in the k th row by the weight $\delta_4^{u_k}$. This is contrast to the WMH-3 model. We will refer to this model as the weighted marginal homogeneity 4 (WMH-4) model.

A special case of the WMH- t ($t=1,2,3,4$) model obtained by putting $\delta_t = 1$ is the MH model and the WMH- t ($t = 1, 2$) model with $\{u_i\}$ replaced by $\{i\}$ is the Tomizawa (1987) DWM models. Note that the OQS model implies the WMH- t ($t = 1, 2, 3, 4$) model.

We obtain the decompositions of the OQS model as follows.

Theorem 3: For t fixed ($t = 1, 2, 3, 4$), the OQS model holds if and only if both the QS and WMH- t models hold.

Proof: For t fixed ($t = 1, 2, 3, 4$), if the OQS model holds, then both the QS and WMH- t models hold.

Conversely, assuming that the QS and WMH- t models hold and then we shall show that the OQS model holds.

First, consider the case of $t = 1$. From the assumption that the QS model holds, we obtain Eq. 1:

$$p_i^-(\delta_1) + p_{ii} + p_i^+ = \mu_i \exp\{v + v_{1(i)}\} \quad (i = 1, \dots, R) \quad (1)$$

Where:

$$\mu_i = \sum_{k=1}^{i-1} \delta_1^{u_i - u_k} \exp\{v_{2(k)} + v_{12(ik)}\} + \sum_{k=i}^R \exp\{v_{2(k)} + v_{12(ki)}\}$$

Also we obtain Eq. 2:

$$p_i^+ + p_{ii} + p_i^-(\delta_1) = \phi_i \xi_i \exp\{v + v_{2(i)}\} \quad (i = 1, \dots, R) \quad (2)$$

Where:

$$\phi_i = \delta_1^{-u_i}, \quad \bar{u} = \frac{1}{R} \sum_{k=1}^R u_k$$

$$\xi_i = \left[\sum_{k=1}^i \exp\{v_{1(k)} + v_{12(ik)}\} + \sum_{k=i+1}^R \delta_1^{u_k - u_i} \exp\{v_{1(k)} + v_{12(ki)}\} \right] / \phi_i$$

Since Eq. 1 is equal to Eq. 2 in terms of the assumption that the WMH-1 model holds, we obtain:

$$\exp\{v_{1(i)}\} = \phi_i h_i \exp\{v_{2(i)}\} \quad (i = 1, \dots, R) \quad (3)$$

where, $h_i = \xi_i / \mu_i$. Substituting Eq. 3 in ξ_i and using $\phi_k / \phi_i = \delta_1^{-u_k}$, yield:

$$f = W f \quad (4)$$

Where:

$$f = (h_1, \dots, h_R)^t$$

and “ t ” denotes the transpose and the (i, m) th element of the $R \times R$ matrix W is:

$$(W)_{im} = \begin{cases} \delta_1^{u_i - u_m} \exp\{v_{2(m)} + v_{12(im)}\} / \mu_i & (i \geq m) \\ \exp\{v_{2(m)} + v_{12(mi)}\} / \mu_i & (i < m) \end{cases}$$

All elements of W are positive and satisfy $WJ_R = J_R$ where $J_R = (1, \dots, 1)^t$ is a vector of order R whose components are all unity. Thus from Eq. 4 and the restriction $\sum_k v_{1(k)} = \sum_k v_{2(k)} = 0$, we get $f = J_R$. Namely we obtain:

$$v_{1(i)} - v_{2(i)} = (\bar{u} - u_i) \log \delta_1 \quad (i = 1, \dots, R)$$

Thus noting that $v_{12(ij)} = v_{12(ji)}$, we see for $i < j$:

$$\begin{aligned} \frac{p_{ij}}{p_{ji}} &= \exp\{(v_{1(i)} - v_{2(i)}) - (v_{1(j)} - v_{2(j)})\} \\ &= \delta_1^{u_j - u_i} \end{aligned}$$

Thus the OQS model holds. The case of $t = 2$ can also be proved in the similar way.

Next consider the case of $t = 3$. From the assumption that the QS model holds, we obtain:

$$p_i^*(\delta_3) = \mu_i^* \exp\{v + v_{1(i)}\} \quad (i = 1, \dots, R) \tag{5}$$

Where:

$$\mu_i^* = \sum_{k=1}^R \delta_3^{u_k} \exp\{v_{2(k)} + v_{12(ik)}\}$$

Also we get:

$$p_i^*(\delta_3) = \delta_3^{u_i - \bar{u}} \xi_i^* \exp\{v + v_{2(i)}\} \quad (i = 1, \dots, R) \tag{6}$$

Where:

$$\xi_i^* = \delta_3^{\bar{u}} \sum_{k=1}^R \exp\{v_{1(k)} + v_{12(ki)}\}$$

Since Eq. 5 is equal to Eq. 6 in terms of the assumption, we obtain:

$$\exp\{v_{1(i)}\} = \delta_3^{u_i - \bar{u}} h_i^* \exp\{v_{2(i)}\} \quad (i = 1, \dots, R) \tag{7}$$

where, $h_i^* = \xi_i^* / \mu_i^*$. Substituting Eq. 7 in ξ_i^* , yield:

$$f^* = W^* f^* \tag{8}$$

Where:

$$f^* = (h_1^*, \dots, h_R^*)^t$$

and the (i, m) th element of the $R \times R$ matrix W^* is:

$$(W^*)_{im} = \delta_3^{u_m} \exp\{v_{2(m)} + v_{12(im)}\} / \mu_i^*$$

with $v_{12(im)} = v_{12(mi)}$. All elements of W^* are positive and satisfy $W^* J_R = J_R$.

Table 2: The values of likelihood ratio chi-squared statistic G^2 for models applied to the data in Table 1

Applied models	Degrees of freedom	G^2	p-value
OQS	9	16.654	0.054
QS	6	6.468	0.373
WMH-1	3	10.464	0.015
WMH-2	3	9.672	0.022
WMH-3	3	9.456	0.024
WMH-4	3	10.663	0.014

Thus from Eq. 8 and the restriction $\sum_k v_{1(k)} = \sum_k v_{2(k)} = 0$, we get $f^* = J_R$. Namely we get:

$$v_{1(i)} - v_{2(i)} = (u_i - \bar{u}) \log \delta_3 \quad (i = 1, \dots, R)$$

Thus the OQS model holds. The case of $t = 4$ can also be proved in the similar way. Thus, the proof is completed.

For a full multinomial sampling, the maximum likelihood estimates of expected frequencies and the parameter δ_t under the WMH-t ($t = 1, 2, 3, 4$) model can be obtained using the Newton-Raphson method to the log-likelihood equations. Each WMH-t model can be tested for goodness-of-fit by likelihood ratio statistic G^2 with $R-2$ degrees of freedom, which is only one less than the number of degrees of freedom for the MH model. The OQS and QS models have $(R+1)(R-2)/2$ and $(R-1)(R-2)/2$, degrees of freedom, respectively.

RESULTS

Consider the data in Table 1 again. Table 2 gives the values of likelihood ratio test statistic G^2 for each model. As described in the section of INTRODUCTION, the OQS and QS models fit these data well. However, according to the test based on the difference between the G^2 values for the OQS and QS models (since the G^2 value is significant at the 0.05 level under the assumption that the QS model holds true), the OQS model does not fit the data well. Also, we see from Table 2 that the WMH-t ($t = 1, 2, 3, 4$) model has a poor fit for the data in Table 1.

DISCUSSION

We see from Theorem 3 that the poor fit of the OQS model is caused by the influence of the lack of structure of the WMH-t ($t = 1, 2, 3, 4$) model rather than the QS model.

CONCLUSION

We have proposed the WMH-t ($t = 1, 2, 3, 4$) models and have given Theorem 3, which is the decompositions of

the OQS model using the new models. Theorem 3 is extensions of Theorems 1 and 2.

Since the WMH-t ($t = 1, 2, 3, 4$) and OQS models are based on the ordered scores, these models and Theorem 3 would be useful for analyzing the square contingency tables with ordered categories with scores $u_1 < \dots < u_R$ (or $u_1 > \dots > u_R$), especially being not equal-interval scores.

Theorem 3 may be useful for exploring the reason for the poor fit when the OQS model fits the data poorly.

REFERENCES

- Agresti, A., 1983. A simple diagonals-parameter symmetry and quasi-symmetry model. *Stat. Probability Lett.*, 1: 313-316. DOI: 10.1016/0167-7152(83)90051-2
- Agresti, A., 2002. *Categorical Data Analysis*. 2nd Edn., John Wiley and Sons, New York, ISBN: 0471360937, pp: 710.
- Bishop, Y.M.M., S.E. Fienberg and P.W. Holland, 1978. *Discrete Multivariate Analysis: Theory and Practice*. 4th Edn., The MIT Press, Cambridge, ISBN: 13: 9780262520409, pp: 557.
- Bowker, A.H., 1948. A test for symmetry in contingency tables. *J. Am. Stat. Assoc.*, 43: 572-574.
- Caussinus, H., 1965. Contribution to the statistical analysis of correlation tables. *Ann. Fac. Sci. Univ. Toulouse*, 29: 77-183.
- Stuart, A., 1955. A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika*, 42: 412-416.
- Tomizawa, S., 1987. Diagonal, weighted marginal homogeneity model and decompositions for linear diagonals-parameter symmetry model. *Commun. Statist. Theory Methods*, 16: 477-488. DOI: 10.1080/03610928708829381