Journal of Statistics Applications & Probability
An International Journal

Decomposition of Symmetry Using Palindromic Symmetry Model in a Two-Way Classification

Kouji Tahata[1], *Kouji Yamamoto*[2] *and Sadao Tomizawa*[1]

[1] Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda City, Japan
[2] Center for Clinical Investigation and Research, Osaka University Hospital, Yamadaoka, Suita, Osaka, Japan

Received: Feb. 14, 2012; Revised Jun. 7, 2012; Accepted Aug. 11, 2012
Published online: 1 Nov. 2012

Abstract: For square contingency tables with ordered categories, we decompose the symmetry model into three models; i.e., the palindromic symmetry, the marginal means equality, and the cumulative subsymmetry models. The palindromic symmetry model is also decomposed into the generalized palindromic symmetry and the extended marginal homogeneity models. The decompositions are applied to the unaided vision data.

Keywords: Decomposition, means equality, ordinal data, palindromic symmetry, square contingency table, symmetry.

1. Introduction

For an $R \times R$ square contingency table with the same row and column classifications having ordered categories, let p_{ij} denote the probability that an observation will fall in the ith row and jth column of the table ($i = 1, \ldots, R; j = 1, \ldots, R$). The symmetry (S) model is defined by

$$p_{ij} = p_{ji} \quad (i \neq j);$$

see Bowker (1948), and Bishop, Fienberg and Holland (1975, p. 282). Caussinus (1965) gave the decomposition of the S model into the quasi-symmetry model and the marginal homogeneity model (although the details of models are omitted). McCullagh (1978) gave the models of asymmetry, namely, the conditional symmetry, the palindromic symmetry (PS), and the generalized palindromic symmetry (GPS) models.

Let

$$G_{ij} = \sum_{s=1}^{i} \sum_{t=j}^{R} p_{st} \quad (i < j),$$

and

$$G_{ij} = \sum_{s=i}^{R} \sum_{t=1}^{j} p_{st} \quad (i > j).$$

The S model may be expressed as

$$G_{ij} = G_{ji} \quad (i \neq j).$$

The PS model is defined by

$$\frac{G_{ij}}{G_{ji}} = \Delta \frac{\gamma_i}{\gamma_{j-1}} \quad (i < j),$$

where we may set, e.g., $\gamma_1 = 1$ without loss of generality. A special case of this model obtained by putting $\Delta = 1$ and $\gamma_1 = \ldots = \gamma_{R-1} = 1$ is the S model. Also, a special case of this model obtained by putting $\gamma_1 = \ldots = \gamma_{R-1} = 1$ is the conditional symmetry model. The PS model with Δ replaced by Δ_i is the GPS model. Note that Tomizawa (1989) gave the decompositions of the conditional symmetry model using the PS model.

If the S model holds, then the PS model holds; however, the converse does not always hold. Thus we are interested in what structure is necessary in addition to the PS model to obtain the S model. The PS model is divided into two structures, as

$$\frac{G_{ij}}{G_{ji}} = \Delta \quad (i < j; j = i + 1),$$

and

$$\frac{G_{ij}}{G_{ji}} = \Delta \frac{\gamma_i}{\gamma_{j-1}} \quad (i < j; j \neq i + 1).$$

* Corresponding author: e-mail: kouji_tahata@is.noda.tus.ac.jp; yamamoto-k@hp-crc.med.osaka-u.ac.jp; tomizawa@is.noda.tus.ac.jp

Therefore we are interested in considering the decomposition of the S model using a model which indicates the structure of $\{G_{ij}\}$ with $|j-i|=1$ and a model which indicates the structure of $\{G_{ij}\}$ with $|j-i|\neq 1$, in addition to the PS model.

The purpose of the present paper is to give such decomposition of the S model using the PS model.

2. Decompositions of the S model

Let X and Y denote the row and the column variables, respectively. Consider the marginal means equality (ME) model as $E(X)=E(Y)$. Let $p_{i\cdot}=\sum_{t=1}^{R} p_{it}$ and $p_{\cdot i}=\sum_{s=1}^{R} p_{si}$ ($i=1,\ldots,R$). We see E(X) = $\sum_{i=1}^{R} ip_{i\cdot}$
$=\sum_{s=1}^{R}\sum_{t=s}^{R} p_{t\cdot}$
$=\sum_{s=1}^{R}\left(1-F_{s-1}^{X}\right)$
$=R-\sum_{i=1}^{R-1} F_i^X$, where $F_i^X=\mathrm{P}(X\leq i)$. Similarly, we see

$$E(Y)=R-\sum_{i=1}^{R-1} F_i^Y,$$

where $F_i^Y=\mathrm{P}(Y\leq i)$. Thus we see E(Y) - E(X) = $\sum_{i=1}^{R-1} F_i^X-\sum_{i=1}^{R-1} F_i^Y$
$=\sum_{i=1}^{R-1} G_{i,i+1}-\sum_{i=1}^{R-1} G_{i+1,i}$.
Therefore the ME model may be expressed as

$$\sum_{i=1}^{R-1} G_{i,i+1}=\sum_{i=1}^{R-1} G_{i+1,i}.$$

This indicates the structure of $\{G_{ij}\}$ with $|j-i|=1$.

On the other hand, as a model which indicates the structure of $\{G_{ij}\}$ with $|j-i|\neq 1$, Tomizawa, Miyamoto and Ouchi (2006) proposed the cumulative subsymmetry (CSS) model as

$$G_{i,i+2}=G_{i+2,i}\quad (i=1,\ldots,R-2).$$

We now obtain the following theorem.
Theorem 1. The S model holds if and only if all the PS, ME and CSS models hold.
Proof. If the S model holds, then all the PS, ME and CSS models hold. Assume that all the PS, ME and CSS models hold, and then we shall show that the S model holds. We see from the PS model that

$$\frac{G_{i,i+1}}{G_{i+1,i}}=\Delta\quad (i=1,\ldots,R-1),$$

and from the ME model that

$$\sum_{i=1}^{R-1} G_{i,i+1}=\sum_{i=1}^{R-1} G_{i+1,i}.$$

Therefore we obtain $\Delta=1$. Thus we see from the PS model that

$$\frac{G_{i,i+2}}{G_{i+2,i}}=\frac{\gamma_i}{\gamma_{i+1}}\quad (i=1,\ldots,R-2).$$

Therefore we see from the CSS model that

$$\gamma_1=\gamma_2=\ldots=\gamma_{R-1}=1.$$

Thus the S model holds. The proof is completed.

By the way, Tomizawa (1984, 1989) considered an extended marginal homogeneity (EMH) model, which is equivalent to

$$\frac{G_{i,i+1}}{G_{i+1,i}}=\Delta\quad (i=1,\ldots,R-1).$$

It is easily seen that the PS model holds if and only if both the GPS model and the EMH model hold. Therefore we also see the following theorem.
Theorem 2. The S model holds if and only if all the GPS, EMH, ME and CSS models hold.

3. Goodness-of-fit test

Let n_{ij} denote the observed frequency in the (i,j)th cell of the $R \times R$ table ($i = 1, \ldots, R; j = 1, \ldots, R$). Assume that a multinomial distribution applies to the $R \times R$ table. Each model can be tested for goodness-of-fit by, e.g., the likelihood ratio chi-squared statistic (denoted by G^2) with the corresponding degrees of freedom. The G^2 is given by

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

where $\hat{m}_{ij}$ is the maximum likelihood estimate of expected frequency m_{ij} under the model. For $\{\hat{m}_{ij}\}$ and the degrees of freedom for each model, see the corresponding literature.

4. An example

Consider the vision data in Table 1. [These data have been analyzed by many statisticians, including Stuart (1955), Bishop et al. (1975, p. 284), McCullagh (1978), Goodman (1979), Tomizawa (1987, 1993), Tomizawa et al. (2006), Tomizawa and Tahata (2007), and Yamamoto and Tomizawa (2012).]

Table 2 gives the values of the likelihood ratio chi-squared statistic G^2 for each model.

Each of the S, GPS and ME models fits the data in Table 1 poorly. Each of the PS, EMH and CSS models fits these data well; especially, the EMH model fits these data very well.

We can see from Theorem 1 that the poor fit of the S model is caused by the influence of the lack of structure of the ME model rather than those of the PS and CSS models. In addition, we can see from Theorem 2 that the poor fit of the S model is caused by the influence of the lack of structure of the GPS and ME models rather than those of the EMH and CSS models. [For the interpretations of the EMH and CSS models applied to these data, see Tomizawa (1987), and Tomizawa et al. (2006), respectively.]

5. Concluding remark

As seen in Example, Theorems 1 and 2 would be useful for seeing the reason for the poor fit of the S model when the S model fits the data poorly. Especially Theorem 2 rather than Theorem 1 is more useful for seeing the reason in the details.

The decompositions obtained in Theorems 1 and 2 should be used to analyze square contingency tables with *ordered* categories because each of decomposed models is not invariant under arbitrary same permutation of the categories of rows and columns.

Table 1

Unaided distance vision of 7477 women aged 30-39 employed in Royal Ordnance factories in Britain from 1943 to 1946; from Stuart (1955).

Right eye grade	Left eye grade				Total
	Best (1)	Second (2)	Third (3)	Worst (4)	
Best (1)	1520	266	124	66	1976
Second (2)	234	1512	432	78	2256
Third (3)	117	362	1772	205	2456
Worst (4)	36	82	179	492	789
Total	1907	2222	2507	841	7477

Table 2

Likelihood ratio chi-square values for models applied to the data in Table 1.

Applied models	Degrees of freedom	Likelihood ratio chi-square
S	6	19.249*
PS	3	6.240
GPS	1	6.180*
EMH	2	0.005
ME	1	11.978*
CSS	2	5.001

* means significant at the 0.05 level.

Acknowledgments

The authors would like to thank the referees for helpful comments.

References

[1] Bishop, Y. M. M., Fienberg, S. E. and Holland, P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, The MIT Press.
[2] Bowker, A. H. (1948). A test for symmetry in contingency tables. *Journal of the American Statistical Association*, **43**, 572-574.
[3] Caussinus, H. (1965). Contribution à l'analyse statistique des tableaux de corrélation. *Annales de la Faculté des Sciences de l'Université de Toulouse*, **29**, 77-182.
[4] Goodman, L. A. (1979). Multiplicative models for square contingency tables with ordered categories. *Biometrika*, **66**, 413-418.
[5] McCullagh, P. (1978). A class of parametric models for the analysis of square contingency tables with ordered categories. *Biometrika*, **65**, 413-418.
[6] Stuart, A. (1955). A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika*, **42**, 412-416.
[7] Tomizawa, S. (1984). Three kinds of decompositions for the conditional symmetry model in a square contingency table. *Journal of the Japan Statistical Society*, **14**, 35-42.
[8] Tomizawa, S. (1987). Diagonal weighted marginal homogeneity models and decompositions for linear diagonals-parameter symmetry model. *Communications in Statistics-Theory and Methods*, **16**, 477-488.
[9] Tomizawa, S. (1989). Decompositions for conditional symmetry model into palindromic symmetry and modified marginal homogeneity models. *Australian Journal of Statistics*, **31**, 287-296.
[10] Tomizawa, S. (1993). Diagonals-parameter symmetry model for cumulative probabilities in square contingency tables with ordered categories. *Biometrics*, **49**, 883-887.
[11] Tomizawa, S., Miyamoto, N. and Ouchi, M. (2006). Decompositions of symmetry model into marginal homogeneity and distance subsymmetry in square contingency tables with ordered categories. *Revstat-Statistical Journal*, **4**, 153-161.
[12] Tomizawa, S. and Tahata, K. (2007). The analysis of symmetry and asymmetry: orthogonality of decomposition of symmetry into quasi-symmetry and marginal symmetry for multi-way tables. *Journal de la Société Française de Statistique*, **148**, 3-36.
[13] Yamamoto, K. and Tomizawa, S. (2012). Analysis of unaided vision data using new decomposition of symmetry. *American Medical Journal*, **3**, 37-42.