

**DECOMPOSITION OF SYMMETRY USING TWO-
RATIOS-PARAMETER SYMMETRY MODEL
AND ORTHOGONALITY FOR SQUARE
CONTINGENCY TABLES**

KOUJI TAHATA and SADA O TOMIZAWA

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

Abstract

Tomizawa [19] proposed the two-ratios-parameter symmetry (TRPS) model which is an extension of the symmetry (S) model. The TRPS model includes the conditional symmetry (CS) model and the linear diagonals-parameter symmetry (LDPS) model in special cases. Caussinus [7] showed that the S model holds if and only if both the quasi-symmetry and marginal homogeneity models hold. Read [14] pointed out that the S model holds if and only if both the CS and global symmetry models hold. Yamamoto et al. [20] showed that the S model holds if and only if both the LDPS and marginal means equality models hold. This paper gives the decompositions of the S model into two or three models using the TRPS model, and considers an orthogonal decomposition such that the goodness-of-fit test statistic for the S model is asymptotically equivalent to the sum of those for the TRPS model and the other model. An example is given.

2000 Mathematics Subject Classification: 62H17.

Keywords and phrases: decomposition, likelihood ratio statistic, linear diagonals-parameter symmetry, orthogonality, separability, square contingency table, symmetry.

Received March 9, 2009

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 i th row and j th column of the table ($i = 1, \dots, r; j = 1, \dots, r$). The symmetry (S) model is defined by

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

see Bowker [5], Caussinus [6, 7] and Bishop et al. ([4], p. 282). This model indicates a structure of symmetry of the probabilities $\{p_{ij}\}$ with respect to the main diagonal of the table.

Tomizawa [19] considered the two-ratios-parameter symmetry (TRPS) model, defined by

$$p_{ij} = \gamma\phi^{j-i}p_{ji} \quad (i < j).$$

This model indicates that the probability that an observation will fall in the (i, j) th cell is $\gamma\phi^{j-i}$ times higher than the probability that it falls in the (j, i) th cell. A special case of this model obtained by putting $\gamma = \phi = 1$ is the S model. Also, special cases of this model obtained by putting $\gamma = 1$ and $\phi = 1$ are the linear diagonals-parameter symmetry (LDPS) model (Agresti, [2]) and the conditional symmetry (CS) model (Read [14]; McCullagh [12]), respectively.

Caussinus [7] considered the quasi-symmetry model and pointed out that the S model holds if and only if both the quasi-symmetry and marginal homogeneity models hold. Consider the global symmetry (GS) model and the marginal means equality (ME) model (see Section 2). Read [14] pointed out that the S model holds if and only if both the CS and GS models hold. Yamamoto et al. [20] showed that the S model holds if and only if both the LDPS and ME models hold. We are now interested in decomposing the S model using the TRPS model (instead of the CS and LDPS models). In the analysis of data, these decompositions for the S model may be useful for seeing the reason for the poor fit when the S model fits the data poorly.

Lang and Agresti [11], and Lang [10] considered the simultaneous modeling of a model for the joint distribution and a model for the marginal distribution. Aitchison [3] discussed the asymptotic separability, which is equivalent to the orthogonality in Read [14] and the independence in Darroch and Silvey [8], of the test statistics for goodness-of-fit of two models (also see Lang and Agresti, [11]; Lang, [10]; Tomizawa [17, 18]; Tomizawa and Tahata, [16]). Thus we are also interested in whether or not the test statistic for the S model is asymptotically equivalent to the sum of the test statistics for decomposed models.

The purpose of this paper is (1) to give the decompositions of the S model using the TRPS model and (2) to prove that the test statistic for the S model is asymptotically equivalent to the sum of those for the TRPS model and the other model. Section 2 presents the decomposition for the S model using the TRPS model. Section 3 shows the orthogonality of test statistics. Section 4 gives an example.

2. Decomposition for Symmetry Using the TRPS Model

For the $r \times r$ table, let X and Y denote the row and column variables, respectively. Consider the GS model defined by

$$\delta_U = \delta_L,$$

where $\delta_U = \sum \sum_{i < j} p_{ij}$ and $\delta_L = \sum \sum_{i > j} p_{ij}$. Note that the S model implies the GS model. Also, consider the ME model defined by

$$\mu_1 = \mu_2,$$

where $\mu_1 = \sum_{i=1}^r i p_{i\cdot}$, $\mu_2 = \sum_{i=1}^r i p_{\cdot i}$, $p_{i\cdot} = \sum_{t=1}^r p_{it}$ and $p_{\cdot i} = \sum_{s=1}^r p_{si}$.

We shall consider a decomposition for the S model. We obtain the following theorem.

Theorem 1. *The S model holds if and only if all the TRPS, GS and ME models hold.*

Proof. If the S model holds, then the TRPS, GS and ME models hold. Assume that all the TRPS, GS and ME models hold, and, then we shall show that the S model holds.

We have

$$\begin{aligned}\mu_1 &= \sum_{i=1}^r ip_i. \\ &= r - \sum_{i=1}^{r-1} F_i^X,\end{aligned}$$

where

$$F_i^X = \sum_{k=1}^i p_k. \quad (i = 1, \dots, r-1).$$

Similarly, we have

$$\mu_2 = r - \sum_{i=1}^{r-1} F_i^Y,$$

where

$$F_i^Y = \sum_{k=1}^i p_{\cdot k} \quad (i = 1, \dots, r-1).$$

Therefore we have

$$\begin{aligned}\mu_2 - \mu_1 &= \sum_{i=1}^{r-1} (F_i^X - F_i^Y) \\ &= \sum_{i=1}^{r-1} (G_{1(i)} - G_{2(i)}),\end{aligned}$$

where

$$G_{1(i)} = \sum_{s=1}^i \sum_{t=i+1}^r p_{st}, \quad G_{2(i)} = \sum_{s=1}^i \sum_{t=i+1}^r p_{ts} \quad (i = 1, \dots, r-1).$$

Since the TRPS model holds, we have

$$p_{st} - p_{ts} = (\gamma\phi^{t-s} - 1)p_{ts} \quad (s < t). \quad (1)$$

Since (1) and the GS model hold, we see

$$\sum_{s < t} (\gamma\phi^{t-s} - 1)p_{ts} = 0,$$

namely,

$$\sum_{k=1}^{r-1} \sum_{s=1}^{r-k} (\gamma\phi^k - 1)p_{s+k,s} = 0. \quad (2)$$

Moreover, since (1) and the ME model hold, we see

$$\sum_{i=1}^{r-1} \left[\sum_{s=1}^i \sum_{t=i+1}^r (\gamma\phi^{t-s} - 1)p_{ts} \right] = 0,$$

namely,

$$\sum_{k=1}^{r-1} \sum_{s=1}^{r-k} k(\gamma\phi^k - 1)p_{s+k,s} = 0. \quad (3)$$

From (2) we see

$$\gamma = \frac{\sum_{l=1}^{r-1} A_l}{\sum_{k=1}^{r-1} A_k \phi^k}, \quad (4)$$

where

$$A_k = \sum_{s=1}^{r-k} p_{s+k,s}.$$

From (3) we see

$$\gamma = \frac{\sum_{l=1}^{r-1} l A_l}{\sum_{k=1}^{r-1} k A_k \phi^k}. \quad (5)$$

From (4) and (5), we obtain

$$\left(\sum_{l=1}^{r-1} A_l \right) \left(\sum_{k=1}^{r-1} k A_k \phi^k \right) - \left(\sum_{l=1}^{r-1} l A_l \right) \left(\sum_{k=1}^{r-1} A_k \phi^k \right) = 0,$$

namely,

$$\sum_{k=1}^{r-1} A_k B_k \phi^{k-1} = 0, \quad (6)$$

where

$$B_k = \sum_{l=1}^{r-1} (k-l) A_l.$$

Noting that

$$\sum_{k=1}^{r-1} A_k B_k = 0,$$

the equation (6) is also expressed as

$$(\phi - 1) \sum_{s=2}^{r-1} (A_{r-1} B_{r-1} + \cdots + A_s B_s) \phi^{s-2} = 0. \quad (7)$$

In addition, for $s = 2, \dots, r-1$,

$$\begin{aligned} A_{r-1} B_{r-1} + \cdots + A_s B_s &= \sum_{t=s}^{r-1} \sum_{l=1}^{r-1} (t-l) A_t A_l \\ &= \sum_{t=s}^{r-1} \sum_{l=1}^{s-1} (t-l) A_t A_l + \sum_{t=s}^{r-1} \sum_{l=s}^{r-1} (t-l) A_t A_l. \end{aligned} \quad (8)$$

The first term on the right-hand side of (8) is positive and the second term equals zero. Thus $A_{r-1} B_{r-1} + \cdots + A_s B_s > 0$ for $s = 2, \dots, r-1$. Therefore, noting that $\phi > 0$, from (7) we obtain $\phi = 1$. Thus, from (2) we obtain $\gamma = 1$. Namely, the S model holds. The proof is completed.

3. Orthogonality of Decomposition for Symmetry

For the decomposition of the S model given in Theorem 1, the orthogonality of test statistics does not hold (see Section 5). Hence, we shall modify Theorem 1 so that the orthogonality of test statistics holds.

Consider a model defined by

$$\delta_U = \delta_L \quad \text{and} \quad \mu_1 = \mu_2. \quad (9)$$

Equation (9) indicates the global symmetry and marginal means equality. We shall call the equation (9) by the GSME model.

From (9) and Theorem 1, we obtain the following lemma.

Lemma 1. *The S model holds if and only if both the TRPS and GSME models hold.*

Assume that the observed frequencies have a multinomial distribution. Let $G^2(M)$ denote the likelihood ratio statistic for testing goodness-of-fit of model M . Table 1 gives the numbers of degrees of freedom (df) for models.

We obtain the following theorem.

Theorem 2. *The following asymptotic equivalence holds:*

$$G^2(S) \simeq G^2(\text{TRPS}) + G^2(\text{GSME}). \quad (10)$$

Proof. The TRPS model is expressed as

$$p_{ij} = \begin{cases} \gamma \phi^{\frac{j-i}{2}} \psi_{ij} & (i < j), \\ \phi^{\frac{j-i}{2}} \psi_{ij} & (i \geq j), \end{cases}$$

where $\psi_{ij} = \psi_{ji}$. Also, the TRPS model can be expressed as

$$\log p_{ij} = \begin{cases} \beta_0 + (j-i)\beta_1 + \Psi_{ij} & (i < j), \\ (j-i)\beta_1 + \Psi_{ij} & (i \geq j), \end{cases} \quad (11)$$

where $\Psi_{ij} = \Psi_{ji}$, $\beta_0 = \log \gamma$, $2\beta_1 = \log \phi$ and $\Psi_{ij} = \log \psi_{ij}$. Let

$$p = (p_{11}, \dots, p_{1r}, p_{21}, \dots, p_{2r}, \dots, p_{r1}, \dots, p_{rr})^t,$$

$$\beta = (\beta_0, \beta_1, \beta_2)^t,$$

where “ t ” denotes the transpose, and where

$$\beta_2 = (\Psi_{11}, \Psi_{12}, \dots, \Psi_{1r}, \Psi_{22}, \Psi_{23}, \dots, \Psi_{2r}, \dots, \Psi_{rr}),$$

is the $1 \times r(r+1)/2$ vector of Ψ_{ij} for $1 \leq i \leq j \leq r$. Then the TRPS model is expressed as

$$\log p = X\beta = (X_0, X_1, X_2)\beta,$$

where X is the $r^2 \times K$ matrix with $K = (r^2 + r + 4)/2$ and

$$X_0 = (\delta_{11}, \dots, \delta_{1r}, \delta_{21}, \dots, \delta_{2r}, \dots, \delta_{r1}, \dots, \delta_{rr})^t; \text{ the } r^2 \times 1 \text{ vector,}$$

$$X_1 = 1_r \otimes J_r - J_r \otimes 1_r; \text{ the } r^2 \times 1 \text{ vector,}$$

and X_2 is the $r^2 \times r(r+1)/2$ matrix of 1 or 0 elements, determined from (11), 1_s is the $s \times 1$ vector of 1 elements and $J_r = (1, 2, \dots, r)^t$, \otimes denotes the Kronecker product, and δ_{ij} is the indicator function, $\delta_{ij} = 1$ if $i < j$, 0 if $i \geq j$. Note that $X_2 1_{r(r+1)/2} = 1_{r^2}$ holds. Note that the matrix X is full column rank which is K . In a similar manner to Haber [9], and Lang and Agresti [11], we denote the linear space spanned by the columns of the matrix X by $S(X)$ with the dimension K . Let U be an $r^2 \times d_1$, where $d_1 = r^2 - K = (r^2 - r - 4)/2$, full column rank matrix such that the linear space spanned by the columns of U , i.e., $S(U)$, is the orthogonal complement of the space $S(X)$. Thus, $U^t X = O_{d_1, K}$, where $O_{d_1, K}$ is the $d_1 \times K$ zero matrix. Therefore the TRPS model is expressed as

$$h_1(p) = O_{d_1},$$

where 0_{d_1} is the $d_1 \times 1$ zero vector and

$$h_1(p) = U^t \log p.$$

The GSME model may be expressed as

$$h_2(p) = 0_{d_2},$$

where $d_2 = 2$ and

$$h_2(p) = Wp,$$

with

$$W = \begin{pmatrix} (-1_{r^2} + 2X_0 + X_{11} + X_{22} + \dots + X_{rr})^t \\ (1_r \otimes J_r - J_r \otimes 1_r)^t \end{pmatrix}; \text{ the } d_2 \times r^2 \text{ matrix,}$$

where $X_{kk}(k = 1, \dots, r)$ is the $r^2 \times 1$ vector, being one of column vectors in X_2 , shouldering Ψ_{kk} . Thus W^t belongs to the space $S(X)$, i.e., $S(W^t) \subset S(X)$. Hence $WU = O_{d_2, d_1}$. From Lemma 1, the S model may be expressed as

$$h_3(p) = 0_{d_3},$$

where $d_3 = d_1 + d_2 = r(r-1)/2$,

$$h_3 = (h_1^t, h_2^t)^t.$$

Note that $h_s(p)$, $s = 1, 2, 3$, are the vectors of order $d_s \times 1$, and d_s , $s = 1, 2, 3$, are the numbers of df for testing goodness-of-fit of the TRPS, GSME and S models, respectively.

Let $H_s(p)$, $s = 1, 2, 3$, denote the $d_s \times r^2$ matrix of partial derivatives of $h_s(p)$ with respect to p , i.e., $H_s(p) = \partial h_s(p) / \partial p^t$. Let $\Sigma(p) = \text{diag}(p) - pp^t$, where $\text{diag}(p)$ denotes a diagonal matrix with i th component of p as i th diagonal component. We see that $H_1(p)p =$

$U^t 1_{r,2} = 0_{d_1}$ since $1_{r,2} \in S(X)$, $H_1(p)diag(p) = U^t$ and $H_2(p) = W$.

Therefore we obtain

$$H_1(p)\Sigma(p)H_2(p)^t = U^tW^t = O_{d_1, d_2}.$$

Thus we obtain $\Delta_3 = \Delta_1 + \Delta_2$ where

$$\Delta_s = h_s(p)^t [H_s(p)\Sigma(p)H_s(p)^t]^{-1} h_s(p). \quad (12)$$

From the asymptotic equivalence of the Wald statistic and likelihood ratio statistic (Rao, [13], Sec. 6e.3; Darroch and Silvey, [8]; Aitchison, [3]) and from (12), we obtain the Equation (10). The proof is completed.

4. An Example

Table 2 taken directly from Agresti ([1], p. 206) is the father's and son's occupational mobility data in Britain. These data have been analyzed by some statisticians including Bishop, et al. ([4], p. 100), Agresti ([1], pp. 205-206) and Yamamoto et al. [20].

Table 3 gives the likelihood ratio chi-square values G^2 for each model. The S model fits the data in Table 2 poorly since $G^2 = 37.46$ with 10 df. Also, the GSME model fits these data poorly, however, the TRPS model fits these data well. Therefore, it is seen from Lemma 1 and Theorem 2 that for these data, the poor fit of the S model is caused by the influence of the lack of structure of the GSME model (rather than the TRPS model). Note that the value of $G^2(S)$ is very close to the value of $G^2(\text{TRPS}) + G^2(\text{GSME})$.

Since the GSME model fits the data in Table 2 poorly, we could state that (i) the probability that the father's status is higher than his son's status is not equal to the probability that the son's status is higher than his father's status, and/or (ii) the mean of father's status is not equal to the mean of son's status. We therefore conclude that there is a significant difference (i) between the "father's status higher" and "son's status higher" groups, and/or (ii) between the father's mean status and son's mean status.

Under the TRPS model the maximum likelihood estimates of γ and ϕ are $\hat{\gamma} = 1.33$ and $\hat{\phi} = 0.96$, and the values of $\hat{\gamma}\hat{\phi}$, $\hat{\gamma}\hat{\phi}^2$, $\hat{\gamma}\hat{\phi}^3$ and $\hat{\gamma}\hat{\phi}^4$ are 1.28, 1.23, 1.18 and 1.13, respectively. The CS model also fits these data well. Under the CS model applied to these data, the maximum likelihood estimates of γ is $\hat{\gamma} = 1.26$. Since the estimated values of γ in the CS model and $\hat{\gamma}\hat{\phi}^i$ ($i = 1, 2, 3, 4$) in the TRPS model are greater than 1, the father's status category is estimated to be less than his son's status category.

5. Concluding Remarks

For Caussinus' [7] decomposition of the S model into Caussinus' [7] quasi-symmetry model and the marginal homogeneity model, Tomizawa and Tahata [16] proved that the test statistic for goodness-of-fit of the S model is asymptotically equivalent to the sum of test statistics for the decomposed models. For Read's [14] decomposition of the S model into the CS and GS models, the test statistic for the S model is equivalent to the sum of those for the CS and GS models. For Yamamoto et al. [20] decomposition of the S model into the LDPS and ME models, Tahata et al. [15] proved that the test statistic for the S model is asymptotically equivalent to the sum of those for decomposed models. In this paper we gave that the test statistic for the S model is asymptotically equivalent to the sum of those for the TRPS and GSME models.

Generally suppose that model M_3 holds if and only if both models M_1 and M_2 hold. As described in Darroch and Silvey [8], (i) when the following asymptotic equivalence holds:

$$G^2(M_3) \simeq G^2(M_1) + G^2(M_2) \quad (13)$$

with $\text{df}(M_3) = \text{df}(M_1) + \text{df}(M_2)$, where $\text{df}(M_k)$ is df for model M_k , if both M_1 and M_2 are accepted (at the α significance level) with high probability, then model M_3 would be accepted; however (ii) when (13) does not hold, such an incompatible situation that both M_1 and M_2 are accepted with high probability but M_3 is rejected with high probability is

quite possible [in fact, Darroch and Silvey [8] showed such an interesting example]. For the orthogonal decompositions of the S model given in Theorem 2, such an incompatible situation would not arise.

The $G^2(S)$ is asymptotically equivalent to $G^2(\text{TRPS}) + G^2(\text{GSME})$ as described by Theorem 2. However, we point out that for the decomposition in Theorem 1, $G^2(S)$ is not asymptotically equivalent to $G^2(\text{TRPS}) + G^2(\text{GS}) + G^2(\text{ME})$ because $G^2(\text{GS}) + G^2(\text{ME})$ is not asymptotically equivalent to $G^2(\text{GSME})$.

Also we point out, for instance, from Theorem 2 that the likelihood ratio statistic for testing goodness-of-fit of the S model assuming that the TRPS model holds true is $G^2(S) - G^2(\text{TRPS})$ and this is asymptotically equivalent to the likelihood ratio statistic for testing goodness-of-fit of the GSME model (i.e., $G^2(\text{GSME})$). Namely, $G^2(\text{GSME})$ can be utilized for testing goodness-of-fit of the GSME model and also for testing goodness-of-fit of the S model assuming that the TRPS model holds true.

References

- [1] A. Agresti, Analysis of Ordinal Categorical Data, Wiley, New York, (1984).
- [2] A. Agresti, A simple diagonals-parameter symmetry and quasi-symmetry model, Stat. Probab. Lett. 1 (1983), 313-316.
- [3] J. Aitchison, Large-sample restricted parametric tests, J. Roy. Statist. Soc. Ser. B 24 (1962), 234-250.
- [4] Y. M. M. Bishop, S. E. Fienberg and P. W. Holland, Discrete Multivariate Analysis: Theory and Practice, The MIT Press, Cambridge, Massachusetts, (1975).
- [5] A. H. Bowker, A test for symmetry in contingency tables, J. Amer. Statist. Assoc. 43 (1948), 572-574.
- [6] H. Caussinus, Some concluding observations, Annales de la Faculté des Sciences de l'Université de Toulouse, Série 6(11) (2002), 587-591.
- [7] H. Caussinus, Contribution à l'analyse statistique des tableaux de corrélation, Annales de la Faculté des Sciences de l'Université de Toulouse, Série 4(29) (1965), 77-182.
- [8] J. N. Darroch and S. D. Silvey, On testing more than one hypothesis, Ann. Math. Stat. 34 (1963), 555-567.

- [9] M. Haber, Maximum likelihood methods for linear and log-linear models in categorical data, *Comput. Stat. Data Anal.* 3 (1985), 1-10.
- [10] J. B. Lang, On the partitioning of goodness-of-fit statistics for multivariate categorical response models, *J. Amer. Statist. Assoc.* 91 (1996), 1017-1023.
- [11] J. B. Lang and A. Agresti, Simultaneously modeling joint and marginal distributions of multivariate categorical responses, *J. Amer. Statist. Assoc.* 89 (1994), 625-632.
- [12] P. McCullagh, A class of parametric models for the analysis of square contingency tables with ordered categories, *Biometrika* 65 (1978), 413-418.
- [13] C. R. Rao, *Linear Statistical Inference and Its Applications*, 2nd edition, Wiley, New York, (1973).
- [14] C. B. Read, Partitioning chi-square in contingency tables: A teaching approach, *Comm. Stat. Theory Methods* 6 (1977), 553-562.
- [15] K. Tahata, H. Yamamoto and S. Tomizawa, Orthogonality of decompositions of symmetry into extended symmetry and marginal equipoment for multi-way tables with ordered categories, *Aust. J. Stat.* 37 (2008), 185-194.
- [16] S. Tomizawa and K. Tahata, 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 (2007), 3-36.
- [17] S. Tomizawa, Orthogonal decomposition of point-symmetry model for two-way contingency tables, *J. Statist. Plann. Inference* 36 (1993), 91-100.
- [18] S. Tomizawa, A decomposition of conditional symmetry model and separability of its test statistic for square contingency tables, *Sankhyā, Ser. B* 54 (1992), 36-41.
- [19] S. Tomizawa, Decompositions for 2-ratios-parameter symmetry model in square contingency tables with ordered categories, *Biom. J.* 29 (1987), 45-55.
- [20] H. Yamamoto, T. Iwashita and S. Tomizawa, Decomposition of symmetry into ordinal quasi-symmetry and marginal equipoment for multi-way tables, *Aust. J. Stat.* 36 (2007), 291-306.

Table 1. Numbers of degrees of freedom for models applied to the $r \times r$ table

Models	Degrees of freedom
S	$r(r - 1) / 2$
CS	$(r + 1)(r - 2) / 2$
LDPS	$(r + 1)(r - 2) / 2$
TRPS	$(r^2 - r - 4) / 2$
GS	1
ME	1
GSME	2

Table 2. Occupational status for British father-son pairs; from Agresti ([1], p. 206)

Father's status	Son's status					Total
	(1)	(2)	(3)	(4)	(5)	
(1)	50	45	8	18	8	129
(2)	28	174	84	154	55	495
(3)	11	78	110	223	96	518
(4)	14	150	185	714	447	1510
(5)	3	42	72	320	411	848
Total	106	489	459	1429	1017	3500

Table 3. Likelihood ratio chi-square values G^2 for models applied to the data in Table 2

Applied models	Degrees of freedom	G^2
S	10	37.46*
CS	9	10.35
LDPS	9	17.13*
TRPS	8	10.02
GS	1	27.12*
ME	1	20.28*
GSME	2	27.44*

* means significant at 5% level.

