

# Orthogonality of Decompositions of Symmetry into Extended Symmetry and Marginal Equipoment for Multi-way Tables with Ordered Categories

Kouji Tahata<sup>1</sup>, Hideharu Yamamoto<sup>2</sup> and Sadao Tomizawa<sup>1</sup>

<sup>1</sup> Dept. of Information Sciences, Tokyo University of Science, Japan

<sup>2</sup> Chugai Clinical Research Center Co., Ltd., Tokyo, Japan

**Abstract:** Yamamoto, Iwashita, and Tomizawa (2007) gave the decompositions of the symmetry (S) model into the extended S and two or three marginal equipoment models for multi-way tables with ordered categories. However the goodness-of-fit test statistic for the S model is not asymptotically equivalent to the sum of those for the decomposed three or four models. This paper gives, for multi-way tables, the modification of their decompositions such that the test statistic for the S model is asymptotically equivalent to the sum of those for the extended S model and one modified marginal equipoment model, which is the constraints combined the marginal means equality and variances equality or combined the marginal means equality, variances equality and correlations equality. Examples are given.

**Zusammenfassung:** Yamamoto et al. (2007) lieferten die Zerlegungen des Symmetrie (S) Modells in das erweiterte S und zwei oder drei marginalen Equipomentmodellen für mehrfach Tafeln mit geordneten Kategorien. Die Teststatistik für die Güte der Anpassung des S Modells ist jedoch nicht asymptotisch äquivalent der Summe jener der zerlegten drei oder vier Modellen. Diese Arbeit gibt für mehrfach Tafeln die Modifizierung derer Zerlegungen so dass die Teststatistik für das S Model asymptotisch äquivalent der Summe jener für das erweiterte S Model und ein modifiziertes marginales Equipoment Modell, was die Beschränkungen darstellt und die Gleichheit der marginalen Erwartungen und die Gleichheit der Varianzen kombiniert oder die Gleichheit der marginalen Erwartungen, Varianzen und Korrelationen kombiniert. Beispiele sind angegeben.

**Keywords:** Likelihood Ratio Statistic, Linear Diagonals-Parameter Symmetry, Ordinal Quasi-Symmetry, Separability, Square Contingency Table.

## 1 Introduction

Consider an  $r^T$  contingency table with ordered categories. Let  $i = (i_1, \dots, i_T)$  for  $i_k = 1, \dots, r$  ( $k = 1, \dots, T$ ), and let  $p_i$  denote the probability that an observation will fall in the  $i$ th cell of the table. Let  $X_k$  ( $k = 1, \dots, T$ ) denote the  $k$ th variable. The symmetry (S- $T$ ) model is defined by

$$p_i = p_j,$$

for any permutation  $j = (j_1, \dots, j_T)$  of  $i = (i_1, \dots, i_T)$  (Bowker, 1948, Agresti, 2002, p. 440).

When  $T = 2$ , Agresti (1983) and Agresti (1984, p. 203) considered the linear diagonals-parameter symmetry (LDPS-2) model. For an  $r^T$  table, Yamamoto et al. (2007) described the LDPS- $T$  model defined by

$$p_i = \left( \prod_{s=1}^T \alpha_s^{i_s} \right) \psi_i,$$

where  $\psi_i = \psi_j$  for any permutation  $j = (j_1, \dots, j_T)$  of  $i = (i_1, \dots, i_T)$ . The LDPS- $T$  model with  $\{\alpha_s = 1\}$  is the S- $T$  model. Note that the LDPS- $T$  model obtained by replacing  $\{i_s\}$  by scores  $\{u_{i_s}\}$  is the ordinal quasi-symmetry model given by Agresti (2002, p. 440).

Consider the marginal mean equality (ME- $T$ ) model defined by  $\mu_1 = \dots = \mu_T$ , where  $\mu_k = E(X_k)$ . Yamamoto et al. (2007) gave the following theorem.

**Theorem 1.** *For the  $r^T$  table ( $T \geq 2$ ), the S- $T$  model holds if and only if both the LDPS- $T$  and ME- $T$  models hold.*

Yamamoto et al. (2007) proposed the extended LDPS- $T$  (denoted by ELDPS- $T$ ) and the more generalized LDPS- $T$  (denoted by GLDPS- $T$ ) models. The GLDPS- $T$  model is defined by

$$p_i = \left( \prod_{s=1}^T \alpha_s^{i_s} \right) \left( \prod_{t=1}^T \beta_t^{i_t^2} \right) \left( \prod_{s=1}^{T-1} \prod_{t=s+1}^T \gamma_{st}^{i_s i_t} \right) \psi_i,$$

where  $\psi_i = \psi_j$  for any permutation  $j = (j_1, \dots, j_T)$  of  $i = (i_1, \dots, i_T)$ . The GLDPS- $T$  model with  $\{\gamma_{st} = 1\}$  is the ELDPS- $T$  model. Note that when  $T = 2$ , the ELDPS-2 model was considered by Tomizawa (1991).

Using the marginal variance equality (VE- $T$ ) and marginal correlation equality (CE- $T$ ) models (see Section 3), Yamamoto et al. (2007) gave the following theorems.

**Theorem 2.** *For the  $r^T$  table ( $T \geq 2$ ), the S- $T$  model holds if and only if all the ELDPS- $T$ , ME- $T$  and VE- $T$  models hold.*

**Theorem 3.** *For the  $r^T$  table ( $T \geq 3$ ), the S- $T$  model holds if and only if all the GLDPS- $T$ , ME- $T$ , VE- $T$  and CE- $T$  models hold.*

Lang and Agresti (1994), and Lang (1996) considered the simultaneous modeling of a model for the joint distribution and a model for the marginal distribution. Aitchison (1962) discussed the asymptotic separability, which is equivalent to the orthogonality in Read (1977) and the independence in Darroch and Silvey (1963), of the test statistics for goodness-of-fit of two models (also see Lang and Agresti, 1994, Lang, 1996, Tomizawa, 1992, 1993, Tomizawa and Tahata, 2007).

We are interested whether or not for Theorems 1, 2 and 3 the test statistic for the S- $T$  model is asymptotically equivalent to the sum of the test statistic for decomposed models.

Theorem 1 is a decomposition for the S- $T$  model into two models, and Theorems 2 and 3 are the decompositions for the S- $T$  model into three or four models. For Theorem 1 the test statistic for the S- $T$  model is asymptotically equivalent to the sum of those for two models (as proved in Section 2), however for Theorems 2 and 3 the test statistic for the

Table 1: Numbers of degrees of freedom (df) for models applied to the  $r^T$  table.

Models	df
S- $T$	$r^T - L$
LDPS- $T$	$r^T - L - (T - 1)$
ELDPS- $T$	$r^T - L - 2(T - 1)$
GLDPS- $T$	$r^T - L - (T^2 + 3T - 6)/2$
ME- $T$	$T - 1$
MV- $T$	$2(T - 1)$
MVC- $T$	$(T^2 + 3T - 6)/2$
Note:	$L = \binom{r+T-1}{T} = \frac{(r+T-1)!}{T!(r-1)!}$

S- $T$  model is not asymptotically equivalent to the sum of those for three or four models (see Section 5 and Appendix). So we are interested in considering the modification of Theorems 2 and 3 such that the test statistic for the S- $T$  model is asymptotically equivalent to the sum of those for the modified two models.

The purpose of this paper is (1) to prove that the test statistic for the S- $T$  model is asymptotically equivalent to the sum of those for the LDPS- $T$  and ME- $T$  models in Section 2, (2) to modify the decomposed three or four models into two models and (3) to consider the modification of Theorems 2 and 3 such that the test statistic for the S- $T$  model is asymptotically equivalent to the sum of those for the modified two models in Section 3.

## 2 Orthogonality of Decomposition for Symmetry

Let  $n_{i_1 \dots i_T}$  denote the observed frequency in the  $(i_1, \dots, i_T)$ th cell of the  $r^T$  table ( $i_k = 1, \dots, r; k = 1, \dots, T$ ) with  $n = \sum \dots \sum n_{i_1 \dots i_T}$ . Assume that the observed frequencies have a multinomial distribution. Let  $G^2(\Omega)$  denote the likelihood ratio statistic for testing goodness-of-fit of model  $\Omega$ . The numbers of degrees of freedom (df) for models are given in Table 1.

First we obtain the following theorem when  $T = 2$ .

**Theorem 4.** *For the  $r \times r$  table, the following asymptotic equivalence holds:*

$$G^2(\text{S-2}) \simeq G^2(\text{LDPS-2}) + G^2(\text{ME-2}).$$

**Proof.** The LDPS-2 model may be expressed as

$$\log p_{ij} = (j - i)\beta_1 + \phi_{ij} \quad i = 1, \dots, r; j = 1, \dots, r, \quad (1)$$

where  $\phi_{ij} = \phi_{ji}$ . Let

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

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

where “ $t$ ” denotes the transpose, and where

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

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

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

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

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

and  $X_2$  is the  $r^2 \times r(r+1)/2$  matrix of 1 or 0 elements, determined from (1),  $1_s$  is the  $s \times 1$  vector of 1 elements and  $J_r = (1, \dots, r)^t$ , and  $\otimes$  denotes the Kronecker product. 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 (1985), and Lang and Agresti (1994), 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  $d_1 = r^2 - K = (r-2)(r+1)/2$ , 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 LDPS-2 model is expressed as

$$h_1(p) = 0_{d_1},$$

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

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

The ME-2 model may be expressed as

$$h_2(p) = 0,$$

where

$$h_2(p) = Wp,$$

with

$$W = (1_r \otimes J_r - J_r \otimes 1_r)^t; \quad \text{the } 1 \times r^2 \text{ vector.}$$

Namely,  $W^t = X_1$ . Thus  $W^t$  belongs to the space  $S(X)$ , i.e.,  $S(W^t) \subset S(X)$ . Hence  $WU = 0_{d_1}^t$ . From Theorem 1, the S-2 model may be expressed as

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

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

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

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 LDPS-2, ME-2 and S-2 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. Let  $\hat{p}$  denote  $p$  with  $\{p_{ij}\}$  replaced by  $\{\hat{p}_{ij}\}$ , where  $\hat{p}_{ij} = n_{ij}/n$ . Then  $\sqrt{n}(\hat{p} - p)$  has asymptotically a normal distribution with mean  $0_{r^2}$  and covariance matrix  $\Sigma(p)$ . Using the delta method (e.g., Agresti, 1984, p. 247),  $\sqrt{n}(h_3(\hat{p}) - h_3(p))$  has asymptotically a normal distribution with mean  $0_{d_3}$  and covariance matrix

$$H_3(p)\Sigma(p)H_3(p)^t = \begin{bmatrix} H_1(p)\Sigma(p)H_1(p)^t & H_1(p)\Sigma(p)H_2(p)^t \\ H_2(p)\Sigma(p)H_1(p)^t & H_2(p)\Sigma(p)H_2(p)^t \end{bmatrix}.$$

We see that  $H_1(p)p = U^t 1_{r^2} = 0_{d_1}$  since  $1_{r^2} \subset S(X)$ ,  $H_1(p)\text{diag}(p) = U^t$  and  $H_2(p) = W$ . Therefore we obtain

$$H_1(p)\Sigma(p)H_2(p)^t = U^t W^t = 0_{d_1}.$$

Thus we obtain  $\Delta_3(\hat{p}) = \Delta_1(\hat{p}) + \Delta_2(\hat{p})$ , where

$$\Delta_s(\hat{p}) = h_s(\hat{p})^t [H_s(\hat{p})\Sigma(\hat{p})H_s(\hat{p})^t]^{-1} h_s(\hat{p}). \tag{2}$$

Under each  $h_s(p) = 0_{d_s}$  ( $s = 1, 2, 3$ ), the Wald statistic  $W_s = n\Delta_s(\hat{p})$  has asymptotically a chi-squared distribution with  $d_s$  degrees of freedom. From (2) we see that  $W_3 = W_1 + W_2$ . From the asymptotic equivalence of the Wald statistic and likelihood ratio statistic (Rao, 1973, Sec. 6e. 3, Darroch and Silvey, 1963; Aitchison, 1962), we obtain Theorem 4. The proof is completed.

Next, we obtain the following theorem for the  $r^T$  table.

**Theorem 5.** *For the  $r^T$  table, the following asymptotic equivalence holds:*

$$G^2(\text{S-T}) \simeq G^2(\text{LDPS-T}) + G^2(\text{ME-T}).$$

The proof of Theorem 5 is omitted because it is obtained in a similar way to the proof of Theorem 4.

### 3 Modified Decomposition Theorems with Orthogonality

For the decomposition for the S-T model into three or four models in Theorems 2 and 3, the orthogonality of test statistics does not hold (see Section 5 and Appendix). Hence, we shall modify the decomposed three or four models into two models as below.

Consider a model defined by

$$\mu_1 = \cdots = \mu_T \quad \text{and} \quad \sigma_1^2 = \cdots = \sigma_T^2, \tag{3}$$

where

$$\mu_k = E(X_k) \quad \text{and} \quad \sigma_k^2 = \text{var}(X_k).$$

Equation (3) is the marginal mean and variance equality (MV-T) model. This is the model combined two constraints of means equality and variances equality. Note that the model defined by only  $\sigma_1^2 = \cdots = \sigma_T^2$ , is the VE-T model in Section 1.

Table 2: Occupational status for British father-son pairs; from Agresti (1984, 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

Next, consider a model defined by

$$\mu_1 = \cdots = \mu_T, \quad \sigma_1^2 = \cdots = \sigma_T^2, \quad \text{and} \quad \rho_{12} = \rho_{13} = \cdots = \rho_{T-1,T}, \quad (4)$$

where

$$\rho_{kl} = \frac{E(X_k X_l) - \mu_k \mu_l}{\sigma_k \sigma_l}.$$

Equation (4) is the marginal mean, variance and correlation equality (MVC- $T$ ) model. This is the model combined three constraints. Note that the model defined by only  $\rho_{12} = \rho_{13} = \cdots = \rho_{T-1,T}$ , is the CE- $T$  model in Section 1.

From (3) and (4), Theorems 2 and 3 may be expressed as the following lemmas.

**Lemma 1.** For the  $r^T$  table ( $T \geq 2$ ), the  $S$ - $T$  model holds if and only if both the ELDPS- $T$  and MV- $T$  models hold.

**Lemma 2.** For the  $r^T$  table ( $T \geq 3$ ), the  $S$ - $T$  model holds if and only if both the GLDPS- $T$  and MVC- $T$  models hold.

Then, in a similar way to the proof of Theorem 4 although the proofs are omitted, we can obtain the following theorems.

**Theorem 6.** For the  $r^T$  table ( $T \geq 2$ ), the following asymptotic equivalence holds:

$$G^2(S-T) \simeq G^2(ELDPS-T) + G^2(MV-T).$$

**Theorem 7.** For the  $r^T$  table ( $T \geq 3$ ), the following asymptotic equivalence holds:

$$G^2(S-T) \simeq G^2(GLDPS-T) + G^2(MVC-T).$$

## 4 Examples

### Example 1

Table 2 taken directly from Agresti (1984, p. 206) is the father's and son's occupational mobility data in Britain. These data have been analyzed by some statisticians including Bishop, Fienberg, and Holland (1975, p. 100), Goodman (1981, 1984), Agresti (1984, pp. 205-206), and Yamamoto et al. (2007).

Table 3: Hemoglobin concentration at baseline, 4 weeks and 8 weeks in carcinomatous anemia patients from a randomized clinical trial; from Yamamoto et al. (2007).

Baseline	4 Weeks	8 Weeks		
		$\geq 10\text{g/dl}$	$8 - 10\text{g/dl}$	$< 8\text{g/dl}$
$\geq 10\text{g/dl}$	$\geq 10\text{g/dl}$	77	7	1
$8 - 10\text{g/dl}$	$\geq 10\text{g/dl}$	43	7	0
$< 8\text{g/dl}$	$\geq 10\text{g/dl}$	3	0	0
$\geq 10\text{g/dl}$	$8 - 10\text{g/dl}$	3	8	1
$8 - 10\text{g/dl}$	$8 - 10\text{g/dl}$	17	16	5
$< 8\text{g/dl}$	$8 - 10\text{g/dl}$	3	8	1
$\geq 10\text{g/dl}$	$< 8\text{g/dl}$	1	1	1
$8 - 10\text{g/dl}$	$< 8\text{g/dl}$	0	2	3
$< 8\text{g/dl}$	$< 8\text{g/dl}$	0	4	3

Table 4a gives the likelihood ratio chi-square values  $G^2$  for each model. The S-2 model fits the data in Table 2 poorly since  $G^2(\text{S-2}) = 37.5$  with 10 df. Also, all of the LDPS-2, ME-2 and MV-2 models fit these data poorly, however, the ELDPS-2 model fits these data well. Therefore, it is seen from Lemma 1 and Theorem 6 that for these data, the poor fit of the S-2 model is caused by the influence of the lack of structure of the MV-2 model (rather than the ELDPS-2 model).

### Example 2

Table 3 taken directly from Yamamoto et al. (2007) is the results of the treatment group only in randomized clinical trials conducted by a pharmaceutical company in anemic patients with cancer receiving chemotherapy. The response is the patient's hemoglobin (Hb) concentration at baseline (before treatment) and following 4 weeks and 8 weeks of treatment. Table 3 shows the  $3 \times 3 \times 3$  array of counts of Hb response that is classified as  $\geq 10\text{g/dl}$ ,  $8 - 10\text{g/dl}$ , and  $< 8\text{g/dl}$ .

The GLDPS-3 model fits the data in Table 3 well, but the other models fit these data poorly (see Table 4b). Therefore, it is seen from Lemma 2 and Theorem 7 that for these data, the poor fit of the S-3 model is caused by the influence of the lack of structure of the MVC-3 model (rather than the GLDPS-3 model).

## 5 Concluding Remarks

Yamamoto et al. (2007) considered decompositions for the symmetry model into three or four models in multi-way tables as Theorems 2 and 3. It may be useful for seeing in more details the reason for the poor fit when the symmetry model fits the data poorly. However, these decompositions do not have such a property that the test statistic for the symmetry model is asymptotically equivalent to the sum of the test statistics for decomposed models. So in this paper we modified these theorems and showed the orthogonality of the test statistic for the decomposition of the symmetry model into the modified two models

Table 4: Likelihood ratio statistics  $G^2$  for models applied to the data in Tables 2 and 3.

(a) For Table 2			(b) For Table 3		
Applied Models	df	$G^2$	Applied Models	df	$G^2$
S-2	10	37.5*	S-3	17	76.2*
LDPS-2	9	17.1*	LDPS-3	15	41.6*
ELDPS-2	8	11.1	ELDPS-3	13	35.5*
			GLDPS-3	11	13.7
ME-2	1	20.3*	ME-3	2	23.8*
MV-2	2	26.3*	MV-3	4	29.2*
			MVC-3	6	53.6*

\*means significant at 5% level

(instead of three or four models).

Generally suppose that model  $\Omega_3$  holds if and only if both models  $\Omega_1$  and  $\Omega_2$  hold. As described in Darroch and Silvey (1963), (i) when the following asymptotic equivalence holds:

$$G^2(\Omega_3) \simeq G^2(\Omega_1) + G^2(\Omega_2) \quad (5)$$

with  $\text{df}(\Omega_3) = \text{df}(\Omega_1) + \text{df}(\Omega_2)$  where  $\text{df}(\Omega_k)$  is the df for model  $\Omega_k$ , if both  $\Omega_1$  and  $\Omega_2$  are accepted (at the  $\alpha$  significance level) with high probability, then  $\Omega_3$  would be accepted; however (ii) when (5) does not hold, such an incompatible situation that both  $\Omega_1$  and  $\Omega_2$  are accepted with high probability but  $\Omega_3$  is rejected with high probability is quite possible (in fact, Darroch and Silvey, 1963 showed such an interesting example). For the orthogonal decompositions of the S- $T$  model given in Theorems 5, 6 and 7 (not in Theorems 2 and 3), such an incompatible situation would not arise.

We point out that, e.g., for the decomposition in Theorem 2,  $G^2(\text{S-}T)$  is not asymptotically equivalent to  $G^2(\text{ELDPS-}T) + G^2(\text{ME-}T) + G^2(\text{VE-}T)$  because  $G^2(\text{ME-}T) + G^2(\text{VE-}T)$  is not asymptotically equivalent to  $G^2(\text{MV-}T)$  (see Appendix), however, the  $G^2(\text{S-}T)$  is asymptotically equivalent to  $G^2(\text{ELDPS-}T) + G^2(\text{MV-}T)$  as described by Theorem 6.

We also point out, e.g., from Theorem 5 that the likelihood ratio statistic for testing goodness-of-fit of the S- $T$  model assuming that the LDPS- $T$  model holds true is  $G^2(\text{S-}T) - G^2(\text{LDPS-}T)$  and this is asymptotically equivalent to the likelihood ratio statistic for testing goodness-of-fit of the ME- $T$  model, i.e.,  $G^2(\text{ME-}T)$ . Namely,  $G^2(\text{ME-}T)$  can be utilized for testing goodness-of-fit of the ME- $T$  model and also for testing goodness-of-fit of the S- $T$  model assuming that the LDPS- $T$  model holds true.

We note that the decomposition considered in this paper should be utilized for analyzing the ordinal data (not the nominal data) because each of models except the S- $T$  model is not invariant under interchanging the arbitrary categories.

Finally we note that the ME- $T$  model may be expressed as  $\mu_1^* = \dots = \mu_T^*$  with  $\mu_k^* = E(g(X_k))$ , where  $g(i) = u_0 + (i - 1)d$ ,  $i = 1, \dots, r$ , and  $u_0$  and  $d$  are specified; namely,  $\{g(i)\}$  may be considered as the known equal-interval scores assigned to the categories. The MV- $T$  and MVC- $T$  models may be expressed using the  $\{g(i)\}$  in the similar way.



## Acknowledgements

The authors would like to thank a referee for the helpful comments.

## Appendix

We shall prove that  $G^2(\text{ME-}T) + G^2(\text{VE-}T)$  is not asymptotically equivalent to  $G^2(\text{MV-}T)$ , thus,  $G^2(\text{S-}T)$  is not asymptotically equivalent to  $G^2(\text{ELDPS-}T) + G^2(\text{ME-}T) + G^2(\text{VE-}T)$ .

**Proof.** We consider the case of  $T = 2$ . The ME-2 and VE-2 models are expressed as  $f_1(p) = 0$ , where

$$f_1(p) = \sum_i \sum_j (j - i)p_{ij},$$

and as  $f_2(p) = 0$ , where

$$f_2(p) = \sum_i \sum_j (j^2 - i^2)p_{ij} - \left( \sum_i \sum_j jp_{ij} \right)^2 + \left( \sum_i \sum_j ip_{ij} \right)^2,$$

respectively. The MV-2 model is expressed as  $f_3(p) = 0$ , where  $f_3(p) = (f_1(p), f_2(p))^t$ . Let  $K_s(p) = \partial f_s(p) / \partial p^t$ ,  $s = 1, 2, 3$ . Using the delta method,  $\sqrt{n}(f_3(\hat{p}) - f_3(p))$  has asymptotically a normal distribution with mean  $(0, 0)^t$  and covariance matrix

$$K_3(p)\Sigma(p)K_3(p)^t = \begin{bmatrix} K_1(p)\Sigma(p)K_1(p)^t & K_1(p)\Sigma(p)K_2(p)^t \\ K_2(p)\Sigma(p)K_1(p)^t & K_2(p)\Sigma(p)K_2(p)^t \end{bmatrix}.$$

Then  $K_1(p)\Sigma(p)K_2(p)^t \neq 0$  although the details are omitted. Therefore, under each  $f_s(p) = 0$  ( $s = 1, 2, 3$ ), the Wald statistic  $W_s = n\Delta_s^*(\hat{p})$  has asymptotically a chi-squared distribution with  $d_s$  degrees of freedom, where

$$\Delta_s^*(\hat{p}) = f_s(\hat{p})^t [K_s(\hat{p})\Sigma(\hat{p})K_s(\hat{p})^t]^{-1} f_s(\hat{p}),$$

and  $d_1 = d_2 = 1$  and  $d_3 = 2$ . Then  $W_3 \neq W_1 + W_2$  since  $K_1(p)\Sigma(p)K_2(p)^t \neq 0$ . Therefore,  $G^2(\text{ME-}2) + G^2(\text{VE-}2)$  is not asymptotically equivalent to  $G^2(\text{MV-}2)$ . In a similar way, the case of  $T \geq 3$  is proven.

## References

- Agresti, A. (1983). A simple diagonals-parameter symmetry and quasi-symmetry model. *Statistics and Probability Letters*, 1, 313-316.
- Agresti, A. (1984). *Analysis of Ordinal Categorical Data*. New York: John Wiley and Sons.
- Agresti, A. (2002). *Categorical Data Analysis* (2nd ed.). New York: John Wiley and Sons.
- Aitchison, J. (1962). Large-sample restricted parametric tests. *Journal of the Royal Statistical Society, Series B*, 24, 234-250.

- Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Massachusetts: The MIT Press.
- Bowker, A. H. (1948). A test for symmetry in contingency tables. *Journal of the American Statistical Association*, 43, 572-574.
- Darroch, J. N., and Silvey, S. D. (1963). On testing more than one hypothesis. *Annals of Mathematical Statistics*, 34, 555-567.
- Goodman, L. A. (1981). Association models and the bivariate normal for contingency tables with ordered categories. *Biometrika*, 68, 347-355.
- Goodman, L. A. (1984). *The Analysis of Cross-Classified Data Having Ordered Categories*. Cambridge: Harvard University Press.
- Haber, M. (1985). Maximum likelihood methods for linear and log-linear models in categorical data. *Computational Statistics and Data Analysis*, 3, 1-10.
- Lang, J. B. (1996). On the partitioning of goodness-of-fit statistics for multivariate categorical response models. *Journal of the American Statistical Association*, 91, 1017-1023.
- Lang, J. B., and Agresti, A. (1994). Simultaneously modeling joint and marginal distributions of multivariate categorical responses. *Journal of the American Statistical Association*, 89, 625-632.
- Rao, C. R. (1973). *Linear Statistical Inference and Its Applications* (2nd ed.). New York: John Wiley and Sons.
- Read, C. B. (1977). Partitioning chi-square in contingency tables: A teaching approach. *Communications in Statistics, Theory and Methods*, 6, 553-562.
- Tomizawa, S. (1991). An extended linear diagonals-parameter symmetry model for square contingency tables with ordered categories. *Metron*, 49, 401-409.
- Tomizawa, S. (1992). A decomposition of conditional symmetry model and separability of its test statistic for square contingency tables. *Sankhyā, Series B*, 54, 36-41.
- Tomizawa, S. (1993). Orthogonal decomposition of point-symmetry model for two-way contingency tables. *Journal of Statistical Planning and Inference*, 36, 91-100.
- 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.
- Yamamoto, H., Iwashita, T., and Tomizawa, S. (2007). Decomposition of symmetry into ordinal quasi-symmetry and marginal equimoment for multi-way tables. *Austrian Journal of Statistics*, 36, 291-306.

#### Authors' Address

Kouji Tahata, Hideharu Yamamoto, and Sadao 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, yamamotohdh@chugai-pharm.co.jp, and tomizawa@is.noda.tus.ac.jp