Journal of Biometrics & Biostatistics
ISSN: 2155-6180

Tahata and Tomizawa, J Biomet Biostat 2011, 2:4
DOI: 10.4172/2155-6180.1000120

Research Article Open Access

Generalized Linear Asymmetry Model and Decomposition of Symmetry for Multiway Contingency Tables

Kouji Tahata* and **Sadao Tomizawa**

Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda City, Chiba, 278-8510, Japan

Abstract

A *k*-order generalization of the linear diagonals-parameter symmetry model is proposed, and related orthogonal decompositions of the generalization are inspected. Applications to randomized clinical trials are given.

Keywords: Linear diagonals-parameter symmetry; Linear ordinal quasi-symmetry; Marginal equimoment; Orthogonal decomposition; Symmetry

Introduction

Consider an r^2 square contingency table with the same row and column classifications. Let p_{ij} denote the probability that an observation will fall in the (i,j) th cell of the table (i=1,…,r;j=1,…,r). For the analysis of square contingency tables, one of our interests is whether or not there is a structure of symmetry (or asymmetry) rather than independence in the table. The symmetry (S^2) model is defined by

$$P_{ij}=\psi_{ij},$$

where $\psi_{ij}=\psi_{ji}$. This indicates that the probability that an observation will fall in the $(i,j)^{th}$ cell is equal to the probability that the observation falls in the $(j,i)^{th}$ cell. As a model that has the weaker restrictions than the S^2 model, Caussinus [1] considered the quasi-symmetry (Q_1^2) model defined by

$$P_{ij}=\mu\alpha_i\beta_j\psi_{ij}$$

where $\psi_{ij}=\psi_{ji}$. A special case of this model with $\{\alpha_i=\beta_i\}$ is the S^2 model. Also Caussinus [1] showed a theorem that the S^2 model holds if and only if both the Q_1^2 and the marginal homogeneity models hold. For the analysis of data, the theorem (say decomposition of the S^2 model) may be useful for seeing the reason for the poor fit when the S^2 model fits the data poorly.

The S^2 and Q_1^2 models indicate the structure of symmetry of cell probabilities and odds-ratios, respectively. As a model that indicates the structure of asymmetry of cell probabilities, Agresti [2] considered the linear diagonals-parameter symmetry (LS^2) model defined by

$$P_{ij}=\mu\alpha^i\beta^j\psi_{ij},$$

where $\psi_{ij}=\psi_{ji}$. This model is a special case of Q_1^2 model. In this way various symmetry and asymmetry models have been proposed by many statisticians (also see Agresti [3]; Tomizawa and Tahata [4]).

Consider an r^T contingency table with ordered categories. Let $i=(i_1,\dots,i_T)$ for $i_k=1,\dots,r$ (k =1,…,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,…,T), denote the k^{th} variable. Tahata et al. [5] considered the linear diagonals-parameter symmetry (LS^T), and extended LS^T (ELS^T) models. The ELS^T model is defined by

$$p_i=\mu\left(\prod_{s=1}^{T}\alpha_{1(s)}^{i_s}\right)\left(\prod_{s=1}^{T}\alpha_{2(s)}^{i_s^2}\right)\psi_i,$$

where $\psi_i=\psi_j$ for any permutation $j=(j_1,\dots,j_T)$ of $i=(i_1,\dots,i_T)$. A special case of this model with $\{\alpha_{2(s)}=1\}$ is the LS^T model. Also a special case of this model with $\{\alpha_{1(s)}=\alpha_{2(s)}=1\}$ is the symmetry (S^T) model (Bhapkar and Darroch [6]; Agresti [3]). Note that when T = 2, the LS^2 model is given by Agresti [2].

Tahata et al. [7] considered the h^{th} linear ordinal quasi-symmetry (LQ_h^T) model (for fixed h (h =1,…,T-1)), defined by

$$p_i=\mu\left(\prod_{k_1=1}^{T}\beta_{k_1}^{i_{k_1}}\right)\left(\prod\prod_{1\le k_1<k_2\le T}\beta_{k_1k_2}^{i_{k_1}i_{k_2}}\right)\times\cdots\times\left(\prod\cdots\prod_{1\le k_1<\cdots<k_h\le T}\beta_{k_1\dots k_h}^{i_{k_1}\cdots i_{k_h}}\right)\psi_i,$$

where $\psi_i=\psi_j$ for any permutation $j=(j_1,\dots,j_T)$ of $i=(i_1,\dots,i_T)$. Note that the LQ_h^T model is a special case of the h^{th} order quasi-symmetry (Q_h^T) model (Bhapkar and Darroch [6]), and LS^T (ELS^T) models are special cases of the first order quasi-symmetry (Q_1^T) model. Note that the Q_h^T model is defined by LQ_h^T with $\left\{\beta_{k_1\dots k_l}^{i_{k_1}\cdots i_{k_l}}\right\}$ replaced by $\left\{\gamma_{k_1\dots k_l(i_{k_1},\dots,i_{k_l})}\right\}, l=1,\dots,h$.

For the analysis of data, when the S^T model does not hold, one may be interested in applying various asymmetry models; for example, the LS^T, ELS^T and LQ_h^T models. If these models do not hold, we are interested in applying a more generalized asymmetry model. In addition we are interested in seeing the reason for the poor fit of the S^T model by using the decomposition of the S^T model. Thus the present paper proposes the generalization of the ELS^T model, and gives the orthogonal decomposition of the S^T model.

Generalized linear asymmetry model

Consider a new model defined by, for a fixed k (k=1,..,r-1),

$$p_i=\mu\left(\prod_{s=1}^{T}\alpha_{1(s)}^{i_s}\right)\left(\prod_{s=1}^{T}\alpha_{2(s)}^{i_s^2}\right)\times\cdots\times\left(\prod_{s=1}^{T}\alpha_{k(s)}^{i_s^k}\right)\psi_i,$$

where $\psi_i=\psi_j$ for any permutation $j=(j_1,\dots,j_T)$ of $i=(i_1,\dots,i_T)$. Without loss

***Corresponding author:** Kouji Tahata, 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, tomizawa@is.noda.tus.ac.jp

Received July 19, 2011; **Accepted** November 07, 2011; **Published** November 15, 2011

Citation: Tahata K, Tomizawa S (2011) Generalized Linear Asymmetry Model and Decomposition of Symmetry for Multiway Contingency Tables. J Biomet Biostat 2:120. doi:10.4172/2155-6180.1000120

Citation: Tahata K, Tomizawa S (2011) Generalized Linear Asymmetry Model and Decomposition of Symmetry for Multiway Contingency Tables. J Biomet Biostat 2:120. doi:10.4172/2155-6180.1000120

of generality, we may set $\alpha_{l(1)}=1$ (l = 1,...,k). We shall refer to this model as the k^{th} linear asymmetry model (denoted by LS_k^T). A special case of the LS_k^T model with $\{\alpha_{1(s)}=\cdots=\alpha_{k(s)}=1\}$, s =1,..,T, is the S^T model. Special cases of the LS_k^T model with $k=1$ and with $k=2$ are LS^T (i.e., LQ_1^T) and ELS^T models, respectively.

The Q_1^T model can be expressed as

$$p_i=\mu\left(\prod_{k_1=1}^{T}\gamma_{k_1(i_{k_1})}\right)\psi_i,$$

where $\psi_i=\psi_j$ for any permutation $j=(j_1,\ldots,j_T)$ of $i=(i_1,\ldots,i_T)$. This model can also be expressed as

$$p_i=\mu\left(\prod_{l=2}^{T}\lambda_{l(i_l)}\right)\varphi_i,$$

where $\varphi_i=\varphi_j$ for any permutation $i=(i_1,\ldots,i_T)$ of $i=(i_1,\ldots,i_T)$ with $\lambda_{l(i_l)}=\gamma_{l(i_l)}/\gamma_{1(i_l)}$ and $\varphi_i=\gamma_{1(i_1)}\ldots\gamma_{1(i_T)}\psi_i$. Then the Q_1^T model with $\{\lambda_{l(i_l)}\}$ replaced by $\{\alpha_{1(l)}^{i_l}\times\cdots\times\alpha_{r-1(l)}^{i_l^{r-1}}\}$ can be expressed as

$$p_i=\mu\left(\prod_{l=2}^{T}\alpha_{1(l)}^{i_l}\right)\times\cdots\times\left(\prod_{l=2}^{T}\alpha_{r-1(l)}^{i_l^{r-1}}\right)\varphi_i,$$

where $\varphi_i=\varphi_j$ for any permutation $j=(j_1,\ldots,j_T)$ of $i=(i_1,\ldots,i_T)$ with Q_1^T model is equivalent to the LS_{r-1}^T model. Therefore we point out that the LS_k^T ($k<r-1$) model is a special case of the Q_1^T model.

Consider the case of $T=2$. For a fixed k (<r), the LS_k^2 model can be expressed as

$$p_{ij}=\mu\alpha_{1(1)}^{i}\alpha_{1(2)}^{j}\times\cdots\times\alpha_{k(1)}^{i^k}\alpha_{k(2)}^{j^k}\psi_{ij},$$

where $\psi_{ij}=\psi_{ji}$. Under this model, the ratio of P_{ij} to P_{ji} is

$$\frac{p_{ij}}{p_{ji}}=\prod_{l=1}^{k}\gamma_l^{j^l-i^l}\quad(i<j),$$

where $\gamma_l=(\alpha_{l(2)}/\alpha_{l(1)})$. Namely, this model indicates that the log ratio of symmetric cells is expressed as the polynomial. Note that the LS_{r-1}^2 model is equivalent to the Q_1^2 model.

Decomposition of symmetry model

For a fixed k ($k=1,..,r$-1), consider a model defined by

$$E(X_1^l)=\cdots=E(X_T^l)\quad(l=1,\ldots,k).$$

We shall refer to this model as the marginal k^{th} moment equality (MME_k^T) model. Then we obtain the following theorem.

Theorem 1: For the r^T table ($T\geq 2$) and k fixed ($k=1,\ldots,r-1$), *the* S^T model holds if and only if both the LS_k^T and MME_k^T models hold.

We give the proof in the Appendix 1. Note that special cases of Theorem 1 with $k=1$ and $k=2$ are given by Tahata et al. [5].

Also although the detail is omitted, we can see that the MME_{r-1}^T model is equivalent to the marginal homogeneity (M_1^T) model defined by

$$P(X_1=i)=\cdots=P(X_T=i)\quad(i=1,\ldots,r),$$

and the LS_{r-1}^T model is equivalent to the Q_1^T model. Namely, a special case of Theorem 1 with $k=r-1$ is identical to the result given by Bhapkar and Darroch [6].

By the way, the MME_1^T model is expressed as

$$E(X_1)=\cdots=E(X_T).$$

Also the MME_2^T model is equivalent to

$$E(X_1)=\cdots=E(X_T),$$

and

$$V(X_1)=\cdots=V(X_T).$$

We are also interested in, for $T\geq3$,

$$E(X_1)=\cdots=E(X_T),$$

and

$$Cov(X_i,X_j)=C\quad(1\leq i<j\leq T),$$

where C is unknown constant. We shall refer to this model as the covariance equality (CE^T) model. Then, in a similar manner to Theorem 1 and Tahata et al. [7], we can obtain the following theorem.

Theorem 2: For the r^T table ($T\geq3$), *the* S^T model holds if and only if both the LQ_2^T and CE^T models hold.

The relationships among models are given in (Figure 1).

Orthogonal decomposition of test statistic

Let $n_{i_1\ldots i_T}$ denote the observed frequency in the $(i_1,\ldots,i_T)^{\text{th}}$ cell of the r^T ($T\geq$ 2) table $(i_k=1,\ldots,r;k=1,\ldots,T)$ with $n=\sum\cdots\sum n_{i_1\ldots i_T}$, and let $m_{i_1\ldots i_T}$ denote the corresponding expected frequency. Assume that $\{n_{i_1\ldots i_T}\}$ have a multinomial distribution. The maximum likelihood estimates of expected frequencies $\{m_{i_1\ldots i_T}\}$ under the LS_k^T, MME_k^T and CE^T models could be obtained using the iterative procedure, for example, the general iterative procedure for log-linear models of Darroch and Ratcliff [8] or using the Newton-Raphson method to the log-likelihood equations.

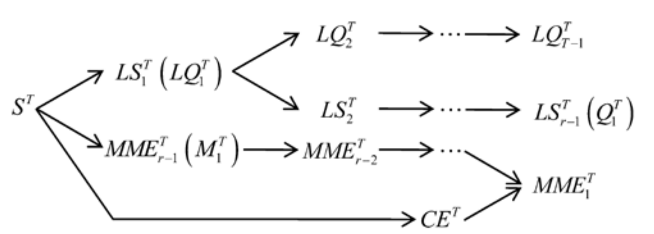

Figure 1: Relationships among models ("A→ B" indicates that model A implies model B).

Citation: Tahata K, Tomizawa S (2011) Generalized Linear Asymmetry Model and Decomposition of Symmetry for Multiway Contingency Tables. J Biomet Biostat 2:120. doi:10.4172/2155-6180.1000120

Each model can be tested for goodness-of-fit by, e.g., the likelihood ratio chi-squared statistic with the corresponding degrees of freedom (df). The numbers of df for the S^T, LS_k^T and MME_k^T models are $r^T - L$, $r^T - L - (T-1)k$ and $(T-1)k$, respectively, where

$$L = \binom{r+T-1}{T} = \frac{(r+T-1)!}{T!(r-1)!}.$$

Also the number of df for the LQ_2^T and CE^T models are $r^T - L - N$ and N, respectively, where $N = (T-2) + T(T-1)/2$.

Let $G^2(M)$ denote the likelihood ratio statistic for testing goodness-of-fit of model M. Thus

$$G^2(M) = 2\sum \cdots \sum n_{i_1 \ldots i_T} \log\left(\frac{n_{i_1 \ldots i_T}}{\hat{m}_{i_1 \ldots i_T}}\right),$$

where $\hat{m}_{i_1 \ldots i_T}$ is the maximum likelihood estimate of expected frequency $m_{i_1 \ldots i_T}$ under model M. Then we obtain the following theorem.

Theorem 3: *For the r^T table ($T \geq 2$) and k fixed $(k = 1, \ldots, r-1)$, the test statistic $G^2(S^T)$ is asymptotically equivalent to the sum of $G^2(LS_k^T)$ and $G^2(MME_k^T)$.*

The proof of Theorem 3 is given in the Appendix 2. In a similar manner to Tahata et al. [5,7], we can obtain the following theorem.

Theorem 4: *For the r^T table ($T \geq 3$), the test statistic $G^2(S^T)$ is asymptotically equivalent to the sum of $G^2(LQ_2^T)$ and $G^2(CE^T)$.*

Note that special cases of Theorem 3 with k =1 and k =2 are given by Tahata et al. [5].

Analysis of data

Analysis of table 1: The data in (Table 1), taken from Stuart [9], are constructed from unaided distance vision of 7477 women aged 30-39 employed in Royal Ordnance factories in Britain from 1943 to 1946 (see, e.g., Caussinus [1]; Tomizawa and Tahata [4]).

Right eye	Left eye grade				
grade	Best (1)	Second (2)	Third (3)	Worst (4)	Total
Best (1)	1520	266	124	66	1976
	(1520.00)	(263.37)	(133.35)	(59.12)	
	(1520.00)	(263.77)	(133.53)	(59.10)	
	(1520.00)	(263.38)	(133.58)	(59.04)	
Second (2)	234	1512	432	78	2256
	(236.63)	(1512.00)	(418.23)	(88.53)	
	(236.23)	(1512.00)	(418.19)	(88.38)	
	(236.62)	(1512.00)	(418.99)	(88.39)	
Third (3)	117	362	1772	205	2456
	(107.65)	(375.77)	(1772.00)	(202.27)	
	(107.47)	(375.81)	(1772.00)	(201.92)	
	(107.42)	(375.01)	(1772.00)	(201.57)	
Worst (4)	36	82	179	492	789
	(42.88)	(71.47)	(181.73)	(492.00)	
	(42.90)	(71.62)	(182.08)	(492.00)	
	(42.96)	(71.61)	(182.43)	(492.00)	
Total	1907	2222	2507	841	7477

Table 1: Unaided distance vision of 7477 women aged 30-39 employed in Royal Ordnance factories in British from 1943 to 1946; from Stuart [9]. (Upper, middle and lower parenthesized values are the maximum likelihood estimates of expected frequencies under the LS_1^2, LS_2^2 and LS_3^2 models, respectively).

The S^2 model fits the data in (Table 1) poorly (see Table 3). By using the decompositions for the S^2 model, we shall consider the reason why the S^2 model fits these data poorly. The LS_k^2 (k=1,2,3) models fit them well, but the MME_k^2 (k =1,2,3) models fit them poorly. So, for example, we see from Theorem 1 with k =1 (i.e., decomposition of the S^2 model into the LS_1^2 and MME_1^2 models) that the poor fit of the S^2 model is caused by the influence of the lack of structure of the MME_1^2 model (rather than the LS_1^2 model). From Theorem 1 with k =2 or k =3, we can obtain similar results.

The LS_3^2 model may be expressed as

$$\frac{p_{ij}}{p_{ji}} = \alpha_{1(2)}^{j-i}\alpha_{2(2)}^{j^2-i^2}\alpha_{3(2)}^{j^3-i^3} \quad (1 \leq i < j \leq 4).$$

Therefore, under this model the probability that a woman's right eye grade is i and her left eye grade is $j(>i)$ is estimated to be $\hat{\alpha}_{1(2)}^{j-i}\hat{\alpha}_{2(2)}^{j^2-i^2}\hat{\alpha}_{3(2)}^{j^3-i^3}$ times higher than the probability that the woman's left eye grade is i and her right eye grade is j, where $\hat{\alpha}_{1(2)} = 1.077$, $\hat{\alpha}_{2(2)} = 1.017$, and $\hat{\alpha}_{3(2)} = 0.998$. Note that $\hat{\alpha}_{l(2)}$ $(l = 1,2,3)$ are the maximum likelihood estimates of $\alpha_{l(2)}$. In this similar manner, the interpretations under the LS_2^2 and LS_1^2 models are obtained although the details are omitted.

According to the test based on the difference between the G^2 values for the LS_2^2 and LS_3^2 models, the LS_2^2 model may be preferable to the LS_3^2 model, and in the similar manner, the LS_1^2 model may be preferable to the LS_2^2 model.

Analysis of table 2: Consider the data in (Table 2) taken from Tahata et al. [5]. These are 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 2) shows the 3×3×3 array of counts of Hb response that is classified as (1) ≥10 g/dl, (2) 8-10 g/dl, and (3) < 8 g/dl. It is reasonable to explore this array for various asymmetries involving time. Namely, we are interested in considering the transition of patient's Hb concentration rather than the interchangeability of evenly spaced points in time with respect to those concentrations. For example, we want to see whether there is an asymmetric transition of those concentrations or not, when the value of those concentration at baseline was given.

We see from (Table 3) that (1) each of the S^3, LS_k^3 $(k = 1,2)$, MME_k^3 (k=1,2,) , and CE^3 models fits the data in (Table 2) poorly, however, (2) the LQ_2^3 model fits them well.

The S^3 model fits the data in (Table 2) poorly (see Table 3). By using the decompositions for the S^3 model, we shall consider the reason why the S^3 model fits these data poorly. The LQ_2^3 model fits them well, but the other models fit them poorly. So, we see from Theorem 2 (i.e., decomposition of the S^3 model into the LQ_2^3 and CE^3 models) that the poor fit of the S^3 model is caused by the influence of the lack of

		8 weeks		
Baseline	4 weeks	(1)	(2)	(3)
(1)	(1)	77	7	1
		(77.00)	(8.53)	(0.22)
(2)	(1)	43	7	0
		(39.80)	(9.35)	(0.13)
(3)	(1)	3	0	0
		(4.72)	(0.60)	(0.02)
(1)	(2)	3	8	1
		(4.67)	(7.69)	(0.74)
(2)	(2)	17	16	5
		(14.96)	(16.00)	(3.30)
(3)	(2)	3	8	1
		(2.80)	(6.42)	(0.70)
(1)	(3)	1	1	1
		(0.07)	(0.41)	(0.59)
(2)	(3)	0	2	3
		(0.33)	(5.28)	(4.03)
(3)	(3)	0	4	3
		(0.39)	(3.27)	(3.00)

The response categories are (1) ≥10 g/dl, (2) 8 — 10 g/dl, (3) < 8 g/dl.

Table 2: Hemoglobin concentration at baseline, 4 weeks and 8 weeks in carcinomatous anemia patients from a randomized clinical trial; from Tahata et al. [5]. (Parenthesized values are the maximum likelihood estimates of expected frequencies under the LQ_2^3 model).

For Table 1			For Table 2		
Models	df	G^2	Models	df	G^2
S^2	6	19.249*	S^3	17	76.186*
LS_1^2	5	7.280	LS_1^3	15	41.550*
LS_2^2	4	7.277	LS_2^3	13	35.466*
LS_3^2	3	7.271	LQ_2^3	13	16.184
MME_1^2	1	11.978*	MME_1^3	2	23.754*
MME_2^2	2	11.982*	MME_2^3	4	29.250*
MME_3^2	3	11.987*	CE^3	4	51.521*

The symbol "*" means significant at 5% level.

Table 3: Likelihood ratio chi-square values G^2 for models applied to the data in (Tables 1 and 2).

structure of the CE^3 model (rather than the LQ_2^3 model).

We shall consider the hypothesis that the LS_1^3 model holds assuming that the LQ_2^3 model holds. Then we can use the test based on the difference between the likelihood ratio chi-squared statistics since the LS_1^3 model is a special case of the LQ_2^3 model. This hypothesis is rejected at the 0.05 significance level since the difference between the two likelihood ratio chi-squared values is 25.366 with 2 df. Therefore the LQ_2^3 model would be preferable to the LS_1^3 model for these data.

Under the LQ_2^3 model, for example, we see

$$\frac{p_{ijk}}{p_{ikj}} = \theta_i^{k-j} \quad (i = 1,2,3; 1 \le j < k \le 3),$$

where $\theta_i = (\beta_3/\beta_2)(\beta_{13}/\beta_{12})^i$. Therefore, under the LQ_2^3 model, the maximum likelihood estimates of $\{\theta_i\}$ are $\hat{\theta}_1 = 1.82$, $\hat{\theta}_2 = 0.63$ and $\hat{\theta}_3 = 0.21$, respectively. Therefore, under the LQ_2^3 model, (i) the conditional probability that the state of the Hb concentration is j at 4 weeks and that is k ($>j$) at 8 weeks, is estimated to be 1.82^{k-j} times higher than the conditional probability that the state of the Hb concentration is k at 4 weeks and that is j at 8 weeks on condition that the patient's Hb concentration is (1) ≥10g/dl at baseline, (ii) those conditional probability is estimated to be 0.63^{k-j} times higher than the corresponding conditional probability on condition that the patient's Hb concentration is (2) 8–10 g/dl at baseline, and (iii) those conditional probability is estimated to be 0.21^{k-j} times higher than the corresponding conditional probability on condition that the patient's Hb concentration is (3) <8 g/dl at baseline. Therefore we could infer that (i) when a patient's Hb concentration is (1) ≥10 g/dl at baseline, those concentration tend to decrease from 4 weeks to 8 weeks since the maximum likelihood estimates of θ_1 is greater than 1, (ii) when a patient's Hb concentration is (2) 8–10 or (3) <8 g/dl at baseline, those concentration tend to increase from 4 weeks to 8 weeks since the maximum likelihood estimates of θ_2 and θ_3 are less than 1.

Concluding Remarks

In this paper, we have proposed the k-order generalization of the linear diagonals-parameter symmetry model that is including the first order quasi-symmetry model, and have given the decomposition of the symmetry model. When the S^T model fits the data poorly, the decomposition of the S^T model (i.e., Theorems 1 and 2) would be useful for seeing the reason for its poor fit. As seen in analysis of (Tables 1, 2), we can see that (1) for the data in (Table 1), the poor fit of the S^2 model is caused by the poor fit of the MME_k^2 models rather than the LS_k^2 ($k=1,2,3$) models, and (2) for the data in (Table 2), the poor fit of the S^3 model is caused by the CE^3 model rather than the LQ_2^3 model.

In Section 3 we have shown a theorem that the S^T model holds if and only if both the LS_k^T and MME_k^T models hold for a fixed k $(k = 1,\ldots, r-1)$. Also, we gave the asymptotic equivalence of test statistic for the S^T model in Theorem 3. Thus, for the orthogonal decomposition of the S^T model into the LS_k^T and MME_k^T models, an incompatible situation, that both the LS_k^T and MME_k^T models are accepted with high probability but the S^T model is rejected with high probability, would not arise. For the orthogonal decomposition of the S^T model into the LQ_2^T and CE^T models, we can also obtain similar results. (For the details of orthogonal decomposition, see Darroch and Silvey [10]).

Acknowledgments

The authors would like to thank a referee to give many helpful comments and suggestions.

References

1. Caussinus H (1965) Contribution à l'analyse statistique des tableaux de corrélation. Annales de la Faculté des Sciences de l'Université de Toulouse 29:

Citation: Tahata K, Tomizawa S (2011) Generalized Linear Asymmetry Model and Decomposition of Symmetry for Multiway Contingency Tables. J Biomet Biostat 2:120. doi:10.4172/2155-6180.1000120

77-183.

2. Agresti A (1983) A simple diagonals-parameter symmetry and quasi-symmetry model. Statistics and Probability Letters 1: 313-316.
3. Agresti A (2002) Categorical Data Analysis. (2nd edn),Wiley, New York.
4. Tomizawa S, 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.
5. Tahata K, Yamamoto H, Tomizawa S (2008) Orthogonality of decompositions of symmetry into extended symmetry and marginal equimoment for multi-way tables with ordered categories. Austrian Journal of Statistics 37: 185-194.
6. Bhapkar VP, Darroch JN (1990) Marginal symmetry and quasi symmetry of general order. Journal of Multivariate Analysis 34: 173-184.
7. Tahata K, Yamamoto H, Tomizawa S (2011) Linear ordinal quasi-symmetry model and decomposition of symmetry for multi-way tables. Mathematical Methods of Statistics 20: 158-164.
8. Darroch JN, Ratcliff D (1972) Generalized iterative scaling for log-linear models. Annals of Mathematical Statistics 43: 1470-1480.
9. Stuart A (1953) The estimation and comparison of strengths of association in contingency tables. Biometrika 40: 105-110.
10. Darroch JN, Silvey SD (1963) On testing more than one hypothesis. Annals of Mathematical Statistics 34: 555-567.

Appendix 1

We shall show the proof of Theorem 1 when T=2. For a fixed k, if the S^2 model holds, then the LS_k^2 and MME_k^2 models hold. Assuming that both the LS_k^2 and MME_k^2 models hold, then we shall show the S^2 model holds. Let $\{p_{ij}^*\}$ denote the cell probabilities which satisfy both the LS_k^2 and MME_k^2 models. Since the LS_k^2 model holds, we see

$$\log p_{ij}^* = \log\mu + \sum_{l=1}^{k}\left(i^l \log\alpha_{l(1)} + j^l \log\alpha_{l(2)}\right) + \log\psi_{ij},$$

where $\psi_{ij} = \psi_{ji}$. Let $\pi_{ij} = \psi_{ij}/c$ with $c = \sum\sum\psi_{ij}$. Then the LS_k^2 and MME_k^2 models are expressed as

$$\log\left(\frac{p_{ij}^*}{\pi_{ij}}\right) = \log\mu c + \sum_{l=1}^{k}\left(i^l \log\alpha_{l(1)} + j^l \log\alpha_{l(2)}\right), \text{ and}$$

$$\mu_1^{l*} = \mu_2^{l*} \quad (l = 1,\ldots,k),$$

where $\mu_1^{l*} = \sum_{s=1}^{r}\sum_{t=1}^{r} s^l p_{st}^*$ and $\mu_2^{l*} = \sum_{s=1}^{r}\sum_{t=1}^{r} s^l p_{ts}^*$. Then we denote $\mu_1^{l*}(=\mu_2^{l*})$ by μ_0^l.

Consider the arbitrary cell probabilities $\{p_{ij}\}$ satisfying

$$\mu_1^l = \mu_2^l = \mu_0^l \quad (l = 1,\ldots,k),$$

where μ_1^l (μ_2^l) denote $\mu_1^{l*}(\mu_2^{l*})$ with $\{p_{ij}^*\}$ replaced by $\{p_{ij}\}$. Let $K(\{a_{ij}\};\{b_{ij}\})$ be the Kullback-Leibler information between $\{a_{ij}\}$ and $\{b_{ij}\}$, where

$$K(\{a_{ij}\};\{b_{ij}\}) = \sum_{i=1}^{r}\sum_{j=1}^{r} a_{ij}\log\left(\frac{a_{ij}}{b_{ij}}\right).$$

From above equations, we see $\sum_{i=1}^{r}\sum_{j=1}^{r}(p_{ij} - p_{ij}^*)\log\left(\frac{p_{ij}^*}{\pi_{ij}}\right) = 0.$

Thus we can obtain $K(\{p_{ij}\};\{\pi_{ij}\}) = K(\{p_{ij}^*\};\{\pi_{ij}\}) + K(\{p_{ij}\};\{p_{ij}^*\})$. Since $\{\pi_{ij}\}$ is fixed, we see $\min_{\{p_{ij}\}} K(\{p_{ij}\};\{\pi_{ij}\}) = K(\{p_{ij}^*\};\{\pi_{ij}\})$, and then $\{p_{ij}^*\}$ uniquely minimize $K(\{p_{ij}\};\{\pi_{ij}\})$.

Let $p_{ij}^{**} = p_{ji}^*$ for $1 \le i, j \le r$. Then noting that $\{\pi_{ij} = \pi_{ji}\}$, we obtain $\min_{\{p_{ij}\}} K(\{p_{ij}\};\{\pi_{ij}\}) = K(\{p_{ij}^{**}\};\{\pi_{ij}\})$, and then $\{p_{ij}^{**}\}$ uniquely minimize $K(\{p_{ij}\};\{\pi_{ij}\})$. Therefore, we see $p_{ij}^* = p_{ji}^*$ for $1 \le i, j \le r$. Namely, the S^2 model holds. The proof is completed in case of $T=2$. For the case of $T > 2$, we can obtain by the same way.

Appendix 2

We shall show the proof of Theorem 3 when $T=2$. For a fixed k, the LS_k^2 model may be expressed as

$$\log p_{ij} = \sum_{l=1}^{k}(j^l - i^l)\beta_l + \phi_{ij} \quad (1 \le i, j \le r),$$

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

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

$$\beta = (\beta_1,\ldots,\beta_k,\phi)^t,$$

where "t" denotes the transpose and

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

is the $1\times r(r+1)/2$ vector of ϕ_{ij} for $1 \le i \le j \le r$. Then the LS_k^2 model is expressed as

$$\log p = X\beta = (X_1,\ldots,X_k,X_{k+1})\beta,$$

where X is the $r^2 \times K$ matrix with $K = r(r+1)/2 + k$ and

$$X_l = 1_r \otimes J_r^l - J_r^l \otimes 1_r; \text{ the } r^2\times 1 \text{ vector} \quad (l = 1,\ldots,k),$$

and X_{k+1} is the $r^2\times r(r+1)/2$ matrix of 1 or 0 elements, determined from the above equation, 1_s is the $s\times 1$ vector of 1 elements and $J_r^l = (1^l, 2^l, \ldots, r^l)^t$ $(l = 1,\ldots,k)$, and $\otimes$ denotes the Kronecker product. Note that $X_{k+1}1_{r(r+1)/2} = 1_{r^2}$ holds. Note that the matrix X is full column rank which is K. We denote the liner 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 $d_1 = r^2 - K = r(r-1)/2 - k$, 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 LS_k^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 MME_k^2 model may be expressed as $h_2(p) = 0$ where $h_2(p) = Wp$ with

$$W = \begin{pmatrix} (1_r \otimes J_r)^t - (J_r \otimes 1_r)^t \\ \vdots \\ (1_r \otimes J_r^k)^t - (J_r^k \otimes 1_r)^t \end{pmatrix}; \text{ the } k\times r^2 \text{ matrix.}$$

Thus W^t belongs to the space $S(X)$, i.e., $S(W^t) \subset S(X)$. Hence $WU = O_{d_2,d_1}$ where $d_2 = k$. 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 $h_3 = (h_1^t, h_2^t)^t$.

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 $\Sigma(p) = diag(p) - pp^t$, where $diag(p)$ 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, $\sqrt{n}(h_3(\hat{p}) - h_3(p))$ 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 $1_{r^2} \subset S(X)$, $H_1(p)diag(p) = U^t$ and $H_2(p) = W$. Therefore we obtain

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

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

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

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 the asymptotic equivalence of the Wald statistic and likelihood ratio statistic, we obtain Theorem 3 when $T = 2$. The proof is completed. The proof of Theorem 3 when $T > 2$ is omitted because it is obtained in the same way.