Open Journal of Statistics, **2013, 3, 9-13**
Published Online December 2013 (http://www.scirp.org/journal/ojs)
http://dx.doi.org/10.4236/ojs.2013.36A002

Scientific Research

Decompositions of Symmetry Using Generalized Linear Diagonals-Parameter Symmetry Model and Orthogonality of Test Statistic for Square Contingency Tables

Kouji Yamamoto[1], Motoki Ohama[2], Sadao Tomizawa[2]
[1]Department of Medical Innovation, Osaka University Hospital, Osaka, Japan
[2]Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Chiba, Japan
Email: yamamoto-k@hp-crc.med.osaka-u.ac.jp, o.motoki1211@gmail.com, tomizawa@is.noda.tus.ac.jp

Received October 10, 2013; revised November 10, 2013; accepted November 17, 2013

ABSTRACT

For square contingency tables with ordered categories, the present paper gives several theorems that the symmetry model holds if and only if the generalized linear diagonals-parameter symmetry model for cell probabilities and for cumulative probabilities and the mean nonequality model of row and column variables hold. It also shows the orthogonality of statistic for testing goodness-of-fit of the symmetry model. An example is given.

Keywords: Cumulative Probability; Global Symmetry; Linear Diagonals-Parameter Symmetry; Mean Equality; Ordinal Category; Orthogonal Test Statistic

1. Introduction

Consider an $R \times R$ square contingency table with the same row and column classifications. Let p_{ij} denote the probability that an observation will fall in the ith row and jth column of the table $(i = 1,\cdots,R; j = 1,\cdots,R)$. Bowker [1] considered the symmetry (S) model defined by

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

This model describes the structure of symmetry with respect to the cell probabilities $\{p_{ij}\}$. As a model which indicates the structure of asymmetry for $\{p_{ij}\}$, Agresti [2] considered the linear diagonals-parameter symmetry (LDPS) model defined by

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

A special case of this model obtained by putting $\delta = 1$ is the S model. Yamamoto and Tomizawa [3] considered the generalized linear diagonals-parameter symmetry (LDPS(K)) model as follows; for a fixed $K\,(K = 0,1,2,\cdots)$,

$$\frac{p_{ij}}{p_{ji}} = \delta^{K+(j-i)} \quad (i < j).$$

Especially the LDPS(0) model is equivalent to the LDPS model.

Let for $i < j$,

$$G_{ij} = \sum_{s=1}^{i}\sum_{t=j}^{R} p_{st} \quad \text{and} \quad G_{ji} = \sum_{s=j}^{R}\sum_{t=1}^{i} p_{st}.$$

The S model may be expressed as

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

Thus the S model also has the structure of symmetry with respect to the cumulative probabilities $\{G_{ij}\}$, $i \neq j$. Miyamoto *et al.* [4] considered the cumulative linear diagonals-parameter symmetry (CLDPS) model defined by

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

which indicates a structure of asymmetry for $\{G_{ij}\}$, $i \neq j$. The CLDPS model is different from the LDPS model. Yamamoto and Tomizawa [3] considered the generalized cumulative linear diagonals-parameter symmetry (CLDPS(*K*)) model as follows; for a fixed $K\left(K=0,1,2,\cdots\right)$,

$$\frac{G_{ij}}{G_{ji}}=\Delta^{K+(j-i)}\quad\left(i<j\right).$$

Especially the CLDPS(0) model is equivalent to the CLDPS model.

Let X and Y denote the row and column variables, respectively. We consider the mean equality (ME) model as

$$E\left(X\right)=E\left(Y\right),$$

where $E\left(X\right)=\sum_{i=1}^{R}ip_{i\cdot}$ and $E\left(Y\right)=\sum_{i=1}^{R}ip_{\cdot i}$, $p_{i\cdot}=\sum_{s=1}^{R}p_{is}$ and $p_{\cdot i}=\sum_{t=1}^{R}p_{ti}$.

Yamamoto *et al.* [5] gave

Theorem 1. *The* S *model holds if and only if both the* LDPS *and* ME *models hold.*

Yamamoto and Tomizawa [6] gave

Theorem 2. *The* S *model holds if and only if both the* CLDPS *and* ME *models hold.*

The present paper gives several decompositions of the S model using the LDPS(*K*) and CLDPS(*K*) models. It also proposes the mean nonequality model, and gives the orthogonal decomposition for testing goodness-of-fit of the S model. An example is given.

2. Decompositions of Symmetry Model

We shall give five kinds of decompositions of the S model using the LDPS(*K*) and CLDPS(*K*) models.

Theorem 3. *For a fixed* $K\left(K=0,1,2,\cdots\right)$, *the* S *model holds if and only if both the* LDPS(*K*) *and* ME *models hold.*

Proof. If the S model holds, then both the LDPS(*K*) and ME models hold. Conversely, assuming that the LDPS(*K*) and ME models hold and then we shall show that the S model holds. The ME model may be expressed as

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

From the LDPS(*K*) model, we see

$$\sum_{i=1}^{R-1}\sum_{s=1}^{i}\sum_{t=i+1}^{R}\delta^{K+(t-s)}p_{ts}=\sum_{i=1}^{R-1}\sum_{s=1}^{i}\sum_{t=i+1}^{R}p_{ts}.$$

Therefore we obtain $\delta=1$. Namely the S model holds. The proof is completed.

Theorem 4. *For a fixed* $K\left(K=0,1,2,\cdots\right)$, *the* S *model holds if and only if both the* CLDPS(*K*) *and* ME *models hold.*

Considering the global symmetry (GS) model as

$$\Pr\left(X<Y\right)=\Pr\left(X>Y\right),$$

namely

$$\sum\sum_{i<j}p_{ij}=\sum\sum_{i<j}p_{ji},$$

we obtain

Theorem 5. *For a fixed* $K\left(K=0,1,2,\cdots\right)$, *the* S *model holds if and only if both the* LDPS(*K*) *and* GS *models hold.*

We shall omit the proofs of Theorems 4 and 5 because these are obtained in a similar manner to the proof of Theorem 3.

For a fixed $K\left(K=0,1,2,\cdots\right)$, consider the mean nonequality (MNE(*K*)) model as follows:

$$E\left(X\right)-E\left(Y\right)=K\left(\Pr\left(X<Y\right)-\Pr\left(X>Y\right)\right),$$

which is

$$\sum\sum_{i<j}\left(K+\left(j-i\right)\right)p_{ij}=\sum\sum_{i<j}\left(K+\left(j-i\right)\right)p_{ji}.$$

This model indicates that the difference between the means of X and Y is K times higher than the difference between the global symmetric probabilities. When $K=0$, the MNE(0) model is identical to the ME model. We obtain

Theorem 6. *For a fixed* $K\left(K=0,1,2,\cdots\right)$, *the* S *model holds if and only if both the* LDPS(*K*) *and* MNE(*K*) *models hold.*

Theorem 7. *For a fixed* $K\left(K=0,1,2,\cdots\right)$, *and for a fixed* $L\left(L=0,1,2,\cdots\right)$, *the* S *model holds if and only if both the* LDPS(*K*) *and* MNE(*L*) *models hold.*

We shall omit the proofs of Theorems 6 and 7 because there are obtained in a similar manner to the proof of Theorem 3. Note that: 1) Theorem 6 is an extension of Theorem 1 because when $K=0$ Theorem 6 is identical to Theorem 1; 2) Theorem 7 is an extension of Theorem 3 because when $L=0$ Theorem 7 is identical to Theorem 3; and 3) Theorem 7 is an extension of Theorem 6 because when $K=L$ Theorem 7 is identical to Theorem 6.

3. Test Statistic and Orthogonality

Let n_{ij} denote the observed frequency in the *i*th row and *j*th column of the $R\times R$ table with $n=\Sigma\Sigma n_{ij}$, and let m_{ij} denote the corresponding expected frequency. Assume that $\left\{n_{ij}\right\}$ has a multinomial distribution. The maximum likelihood estimates of expected frequencies $\left\{m_{ij}\right\}$ under each model could be obtained, for example, using the Newton-Raphson method to the log-likelihood equations. Each model (say, model M) can be tested for goodness-of-fit by the likelihood ratio chi-squared statistic $G^2\left(M\right)$ with the corresponding degrees of freedom, defined by

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

where $\hat{m}_{ij}$ is the maximum likelihood estimate of m_{ij} under the model. The number of degrees of freedom for the S model is $R(R-1)/2$, and that for each of the LDPS(*K*) and CLDPS(*K*) models is $(R-2)(R+1)/2$ (being one less than that for the S model). That for each of ME, GS, and MNE(*K*) models is 1. Note that the number of degrees of freedom for the S model is equal to the sum of those for the decomposed models.

Lang and Agresti [7] and Lang [8] considered the simultaneous modeling of a model for the joint distribution and a model for the marginal distribution. Aitchison [9] discussed the asymptotic separability, which is equivalent to the orthogonality in Read [10] and the independence in Darroch and Silvey [11], of the test statistic for goodness-of-fit of two models (also see Tomizawa and Tahata [12], Tahata *et al.* [13], and Tahata and Tomizawa [14]). On the orthogonality of test statistic for models in Theorem 6, we obtain.

Theorem 8. *For a fixed* $K(K=0,1,2,\cdots)$, *test statistic* $G^2(\mathrm{S})$ *is asymptotically equivalent to the sum of* $G^2(\mathrm{LDPS}(K))$ *and* $G^2(\mathrm{MNE}(K))$.

Proof. The LDPS(*K*) model may be expressed as

$$\log p_{ij} = \begin{cases} (K+j-i)\beta_1 + \phi_{ij} & (i<j), \\ \phi_{ii} & (i=j), \\ -(K+i-j)\beta_1 + \phi_{ij} & (i>j), \end{cases} \quad (1)$$

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

$$p = (p_{11},\cdots,p_{1R},p_{21},\cdots,p_{2R},\cdots,p_{R1},\cdots,p_{RR})^t,$$

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

where "*t*" denotes the transpose, and

$$\beta_2 = (\phi_{11},\phi_{12},\cdots,\phi_{1R},\phi_{22},\cdots,\phi_{2R},\cdots,\phi_{R-1,R},\phi_{RR})$$

is the $1\times R(R+1)/2$ vector. The LDPS**(*K*)** model is expressed as

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

where X is the $R^2 \times L$ matrix with $L = (R^2+R+2)/2$, and X_1 is the $R^2\times 1$ vector with

$$X_1 = (x_{11},\cdots,x_{1R},x_{21},\cdots,x_{2R},\cdots,x_{R1},\cdots,x_{RR})^t,$$

where

$$x_{ij} = \begin{cases} K+j-i & (i<j), \\ 0 & (i=j), \\ -(K+i-j) & (i>j), \end{cases}$$

and X_2 is $R^2 \times R(R+1)/2$ matrix of 0 or 1 elements determined from (1). The matrix X is full column rank which is L. In a similar manner to Haber [15], Lang and Agresti [7], and Tahata and Tomizawa [16], we denote the linear space spanned by columns of the matrix X by $S(X)$ with the dimension L. Note that $X_2 1_{R(R+1)/2} = 1_{R^2}$ where 1_t is the $t\times 1$ vector of 1 elements, and thus $1_{R^2} \in S(X)$. Let U be an $R^2 \times d_1$, where $d_1 = R^2 - L = (R-2)(R+1)/2$, full column rank matrix such that the linear space $S(U)$ is the orthogonal component of the space $S(X)$. Thus, $U^t X = O_{d_1,L}$, where O_{st} is the $s\times t$ zero matrix. Therefore, the LDPS(*K*) model is expressed as

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

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

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

The MNE(*K*) model may be expressed as

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

where $d_2 = 1$,

$$h_2(p) = X_1^t p.$$

Note that $X_1^t U = 0_{d_1}$. From Theorem 6, the S model may be expressed as

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

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

$$h_3 = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}.$$

Note that $d_s\ (s=1,2,3)$ are the numbers of degrees of freedom for testing goodness-of-fit of the LDPS(*K*), MNE(*K*) 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 $\Sigma(p) = diag(p) - pp^t$, where $diag(p)$ denotes a diagonal matrix with *i*th component of p as *i*th diagonal component. We see that

$$H_1(p)p = U^t 1_{R^2} = 0_{d_1},$$

because $1_{R^2} \in S(X)$, and that

$$H_1(p)diag(p) = U^t,$$

$$H_2(p) = X_1^t.$$

Thus we obtain

$$H_1(p)\Sigma(p)H_2(p)^t = U^t X_1 = 0_{d_1}.$$

Therefore we obtain $\Delta_3 = \Delta_1 + \Delta_2$, where

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

From the asymptotic equivalence of the Wald statistic and the likelihood ratio statistic (Rao [17], Darroch and Silvey [11], Aitchison [9]), we obtain Theorem 8. The proof is completed.

4. Analysis of Data

Table 1 taken directly from Agresti [18, p. 232] summarizes responses to the questions "How successful is the government in (1) providing health care for the sick? (2) Protecting the environment?".

Table 2 gives the values of the likelihood ratio test statistic G^2 for models applied to these data. The S model does not fit these data so well. Also, each of the ME (*i.e.*, MNE(0)), MNE(*K*) $(K = 1, 2, \cdots, 5)$ and the GS models does not fit these data so well. However each of the LDPS(*K*) models $(K = 0, 1, \cdots, 5)$ and the CLDPS(*K*) models $(K = 1, 2, \cdots, 5)$ fit these data very well. Using Theorems 3 through 7 (including Theorems 1 and 2), we shall consider the reason why the S model fits these data poorly. For the structure of cell probabilities $\{p_{ij}\}$, we see from Theorems 3, 5, 6 and 7 that the poor fit of the S model is caused by the influence of the lack of structure of the ME model (the GS model or the MNE(*K*) model $(K = 1, 2, \cdots, 5)$) rather than the LDPS(*K*) model $(K = 0, 1, \cdots, 5)$. For the structure of cumulative probabilities $\{G_{ij}\}$, $i \neq j$, we see from Theorem 4 that the poor fit of the S model is caused by the influence of the lack of structure of the ME model rather than the CLDPS(*K*) model $(K = 1, 2, \cdots, 5)$.

Table 1. Data on success of US government in providing health care and protecting the environment; from Agresti [18, p. 232]. (Upper and lower parenthesized values are maximum likelihood estimates of expected frequencies under the LDPS(1) and CLDPS(4) models, respectively.)

Health Care	Environment			Total
	Successful	Mixed	Unsuccessful	
Successful	199	81	83	363
	(199.00)	(82.04)	(83.79)	
	(199.00)	(83.00)	(83.97)	
Mixed	129	167	112	408
	(127.96)	(167.00)	(109.78)	
	(127.24)	(167.00)	(108.73)	
Unsuccessful	164	169	363	696
	(163.21)	(171.22)	(363.00)	
	(163.09)	(171.97)	(363.00)	
Total	492	417	558	1467

Table 2. Values of likelihood ratio chi-squared statistic G^2 for models applied to the data in Table 1.

Models	Degrees of freedom	G^2	*p*-value
S	3	49.77*	<0.001
ME (*i.e.*, MNE(0))	1	48.82*	<0.001
MNE(1)	1	49.68*	<0.001
MNE(2)	1	49.43*	<0.001
MNE(3)	1	49.14*	<0.001
MNE(4)	1	48.90*	<0.001
MNE(5)	1	48.71*	<0.001
GS	1	47.39*	<0.001
LDPS(0)	2	0.76	0.685
LDPS(1)	2	0.11	0.948
LDPS(2)	2	0.40	0.821
LDPS(3)	2	0.69	0.708
LDPS(4)	2	0.93	0.630
LDPS(5)	2	1.11	0.575
CLDPS(0)	2	19.56*	<0.001
CLDPS(1)	2	3.82	0.148
CLDPS(2)	2	0.87	0.647
CLDPS(3)	2	0.28	0.870
CLDPS(4)	2	0.24	0.888
CLDPS(5)	2	0.35	0.841

*Means significant at the 0.05 level.

5. Concluding Remarks

We have given the new five kinds of decompositions of the S model. Theorems 3 and 4 are extensions of Theorems 1 and 2, respectively. Theorem 5 is another decomposition of the S model. Theorem 6 is another extension of Theorem 1, and Theorem 7 is an extension of Theorem 3. These theorems may be useful for seeing in more details the reason for the poor fit when the S model fits the data poorly.

From the orthogonality of test statistic given by Theorem 8, we point out that for instance, the likelihood ratio chi-squared statistic for testing goodness-of-fit of the S model assuming that the LDPS(*K*) model holds true is $G^2(\text{S}) - G^2(\text{LDPS}(K))$ and this is asymptotically equivalent to the likelihood ratio chi-squared statistic for testing goodness-of-fit of the MNE(*K*) model, *i.e.*, $G^2(\text{MNE}(K))$. We see that for the data in **Table 1** the value of $G^2(\text{S})$ is very close to the sum of the values of $G^2(\text{LDPS}(K))$ and $G^2(\text{MNE}(K))$ (see **Table 2**). The orthogonal decomposition of the S model into the

LDPS(K) and MNE(K) models would guarantee that (1) if both the LDPS(K) and MNE(K) models are accepted (e.g., at the 0.05 significance level) with high probability, then the S model would be accepted, and (2) it would be impossible to arise such an incompatible situation that both the LDPS(K) and MNE(K) models are accepted with high probability but the S model is rejected with high probability. Therefore, in particular Theorems 6 and 8 would be useful for analyzing the data.

REFERENCES

[1] A. H. Bowker, "A Test for Symmetry in Contingency Tables," *Journal of the American Statistical Association*, Vol. 43, No. 244, 1948, pp. 572-574. http://dx.doi.org/10.1080/01621459.1948.10483284

[2] A. Agresti, "A Simple Diagonals-Parameter Symmetry and Quasi-Symmetry Model," *Statistics and Probability Letters*, Vol. 1, No. 6, 1983, pp. 313-316. http://dx.doi.org/10.1016/0167-7152(83)90051-2

[3] K. Yamamoto and S. Tomizawa, "Statistical Analysis of Case-Control Data of Endometrial Cancer Based on New Asymmetry Models," *Journal of Biometrics and Biostatistics*, Vol. 3, No. 5, 2012, pp. 1-4. http://dx.doi.org/10.4172/2155-6180.1000147

[4] N. Miyamoto, W. Ohtsuka and S. Tomizawa, "Linear Diagonals-Parameter Symmetry and Quasi-Symmetry Models for Cumulative Probabilities in Square Contingency Tables with Ordered Categories," *Biometrical Journal*, Vol. 46, No. 6, 2004, pp. 664-674. http://dx.doi.org/10.1002/bimj.200410066

[5] H. Yamamoto, T. Iwashita and S. Tomizawa, "Decomposition of Symmetry into Ordinal Quasi-Symmetry and Marginal Equimoment for Multi-way Tables," *Austrian Journal of Statistics*, Vol. 36, No. 4, 2007, pp. 291-306.

[6] K. Yamamoto and S. Tomizawa, "Analysis of Unaided Vision Data Using New Decomposition of Symmetry," *American Medical Journal*, Vol. 3, No. 1, 2012, pp. 37-42. http://dx.doi.org/10.3844/amjsp.2012.37.42

[7] J. B. Lang and A. Agresti, "Simultaneously Modeling Joint and Marginal Distributions of Multivariate Categorical Responses," *Journal of the American Statistical Association*, Vol. 89, No. 426, 1994, pp. 625-632. http://dx.doi.org/10.1080/01621459.1994.10476787

[8] J. B. Lang, "On the Partitioning of Goodness-of-Fit Statistics for Multivariate Categorical Response Models," *Journal of the American Statistical Association*, Vol. 91, No. 435, 1996, pp. 1017-1023. http://dx.doi.org/10.1080/01621459.1996.10476972

[9] J. Aitchison, "Large-Sample Restricted Parametric Tests," *Journal of the Royal Statistical Society*: *Series B*, Vol. 24, No. 1, 1962, pp. 234-250.

[10] C. B. Read, "Partitioning Chi-Square in Contingency Tables: A Teaching Approach," *Communications in Statistics-Theory and Methods*, Vol. 6, No. 6, 1977, pp. 553-562. http://dx.doi.org/10.1080/03610927708827513

[11] J. N. Darroch and S. D. Silvey, "On Testing More than One Hypothesis," *Annals of Mathematical Statistics*, Vol. 34, No. 2, 1963, pp. 555-567. http://dx.doi.org/10.1214/aoms/1177704168

[12] 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 Francaise de Statistique*, Vol. 148, No. 3, 2007, pp. 3-36.

[13] K. Tahata, H. Yamamoto and S. Tomizawa, "Orthogonality of Decompositions of Symmetry into Extended Symmetry and Marginal Equimoment for Multi-Way Tables with Ordered Categories," *Austrian Journal of Statistics*, Vol. 37, No. 2, 2008, pp. 185-194.

[14] K. Tahata and S. Tomizawa, "Orthogonal Decomposition of Point-Symmetry for Multiway Tables," *Advances in Statistical Analysis*, Vol. 92, No. 3, 2008, pp. 255-269. http://dx.doi.org/10.1007/s10182-008-0070-5

[15] M. Haber, "Maximum Likelihood Methods for Linear and Log-Linear Models in Categorical Data," *Computational Statistics and Data Analysis*, Vol. 3, No. 1, 1985, pp. 1-10. http://dx.doi.org/10.1016/0167-9473(85)90053-2

[16] K. Tahata and S. Tomizawa, "Double Linear Diagonals-Parameter Symmetry and Decomposition of Double Symmetry for Square Tables," *Statistical Methods and Applications*, Vol. 19, No. 3, 2010, pp. 307-318. http://dx.doi.org/10.1007/s10260-009-0127-y

[17] C. R. Rao, "Linear Statistical Inference and its Applications," 2nd Edition, John Wiley, New York, 1973. http://dx.doi.org/10.1002/9780470316436

[18] A. Agresti, "Analysis of Ordinal Categorical Data," 2nd Edition, John Wiley, Hoboken, 2010. http://dx.doi.org/10.1002/9780470594001