Journal of Modern Applied Statistical Methods

Volume 9 | Issue 1 Article 25

5-1-2010

Symmetry Plus Quasi Uniform Association Model and Its Orthogonal Decomposition for Square Contingency Tables

Kouji Yamamoto
Osaka University Hospital, Suita City, Japan, yamamoto-k@hp-crc.med.osakau.ac.jp

Sadao Tomizawa
Tokyo University of Science, Noda City, Japan, tomizawa@is.noda.tus.ac.jp

Journal of Modern Applied Statistical Methods
May 2010, Vol. 9, No. 1, 255-262

1538 – 9472/10/$95.00

Symmetry Plus Quasi Uniform Association Model and Its Orthogonal Decomposition for Square Contingency Tables

Kouji Yamamoto
Osaka University Hospital,
Suita City, Japan

Sadao Tomizawa
Tokyo University of Science,
Noda City, Japan

A model is proposed having the structure of both symmetry and quasi-uniform association (SQU model) and provides a decomposition of the SQU model. It is also shown with examples that the test statistic for goodness-of-fit of the SQU model is asymptotically equivalent to the sum of those for the decomposed models.

Key words: Likelihood ratio statistic, marginal homogeneity, quasi-symmetry, quasi-uniform association, separability.

Introduction

For the $r \times r$ square contingency table, let p_{ij} denote the probability that an observation will fall in the i^{th} row and j^{th} column of the table ($i = 1,\ldots,r; j = 1,\ldots,r$). For the analysis of two-way contingency tables with ordered categories, Goodman (1979) considered some association models, for example, the uniform association model, which is a generalization of the independence model. Goodman (1979) also observed that regular multiplicative models for ordinal variables fit square contingency tables well when the cells on the main diagonal are ignored, thus, he proposed the quasi-uniform association (QU) model, defined by

$$p_{ij} = \begin{cases} \mu\alpha_i\beta_j\theta^{ij} & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases}$$

when $\theta = 1$, this model is the quasi-independence (QI) model (Bishop, Fienberg & Holland, 1975, p. 178).

The symmetry (S) model considered by Bowker (1948) is defined by

$$p_{ij} = \psi_{ij} \quad (i = 1,\ldots,r; j = 1,\ldots,r),$$

where $\psi_{ij} = \psi_{ji}$ (Bishop, et al., 1975, p. 282). This model describes a structure of symmetry of the cell probabilities $\{p_{ij}\}$ with respect to the main diagonal of the table. The quasi-symmetry (QS) model considered by Caussinus (1965) is defined by

$$p_{ij} = \mu\alpha_i\beta_j\psi_{ij} \quad (i = 1,\ldots,r; j = 1,\ldots,r),$$

where $\psi_{ij} = \psi_{ji}$. The odds ratio for rows i and j $(> i)$, and columns s and t $(> s)$ are denoted by $\theta_{(ij;st)}$; thus,

$$\theta_{(ij;st)} = \frac{(p_{is}p_{jt})}{(p_{js}p_{it})}$$

Using odds ratios, the QS model may be expressed as

$$\theta_{(ij;st)} = \theta_{(st;ij)} \quad (i < j; s < t).$$

Kouji Yamamoto is an Assistant Professor on Center for Clinical Investigation and Research. Email him at: yamamoto-k@hp-crc.med.osaka-u.ac.jp. Sadao Tomizawa is a Professor on the Faculty of Science and Technology in the Department of Information Sciences. Email him at: tomizawa@is.noda.tus.ac.jp.

Therefore this model indicates a structure of symmetry with respect to the odds ratios. A special case of this model obtained by putting $\{\alpha_i = \beta_i\}$ is the S model. Also, each of the QI and QU models is a special case of the QS model. The marginal homogeneity (MH) model is defined by

$$p_{i\cdot} = p_{\cdot i} \quad (i = 1, \ldots, r),$$

where $p_{i\cdot} = \sum_{t=1}^{r} p_{it}$ and $p_{\cdot i} = \sum_{s=1}^{r} p_{si}$ (Stuart, 1955).

Decomposition of the S Model (Caussinus, 1965)

Theorem 1

The S model holds if and only if both the QS and MH models hold.

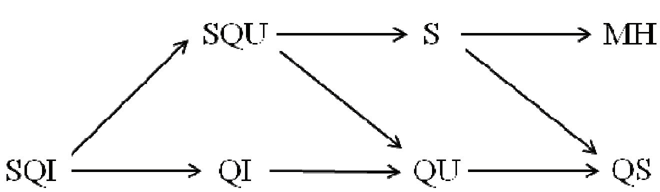
The symmetry plus quasi-independence (SQI) model introduced by Goodman (1985) is defined by

$$p_{ij} = \begin{cases} \mu\alpha_i\alpha_j & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases}$$

This model is a special case of the S model obtained by substituting $\{\psi_{ij} = \alpha_i\alpha_j\}$ for $i \neq j$.

The purpose of this study is to: (1) propose a model that can be used to simultaneously study both symmetry and quasi-uniform association (the SQU model), (2) provide decomposition of the proposed model, and (3) show the orthogonality of decomposition with respect to the goodness-of-fit test statistic.

Figure 1. Relationships among the Models



Proposed Model

Consider a model defined by:

$$p_{ij} = \begin{cases} \mu\alpha_i\alpha_j\theta^{ij} & (i \neq j), \\ \psi_{ii} & (i = j). \end{cases}$$

This model indicates that both the S and QU models hold simultaneously. Thus, this model shall be referred to as the symmetry plus quasi-uniform association (SQU) model. The SQU model is an extension of the SQI model. Under the SQU model, the row marginal distribution is identical with the column marginal distribution. Using odds ratios, the SQU model may be expressed as

$$\theta_{(ij;st)} = \theta^{(j-i)(t-s)} \quad (i \neq s, i \neq t, j \neq s, j \neq t).$$

This model has uniform local association for cells off the main diagonal of the table. Figure 1 shows the relationships among the models.

Decompositions of the Models

Theorem 2

The SQU model holds if and only if both the QU and MH models hold.

Theorem 2 Proof

If the SQU model holds, then the QU and MH models hold. Conversely, if both the QU and MH models hold, then the QS model holds. Therefore, from Theorem 1 it may be stated that the S model holds. Thus, SQU model holds. The proof is completed and the following corollary is obtained because the SQI model is a special case of the SQU model with $\theta = 1$.

Corollary 1

The SQI model holds if and only if both the QI and MH models hold.

Orthogonality of Decomposition of Test Statistic for the Models

Let n_{ij} denote the observed frequency in the $(i, j)^{\text{th}}$ cell of the table $(i = 1, \ldots, r; j = 1, \ldots, r)$. Assume that a multinomial distribution applies to the $r \times r$ table. The maximum likelihood estimates of expected frequencies under the models described in this paper could be obtained using an iterative procedure, for example, the general iterative procedure for log-linear models of Darroch and Ratcliff (1972) or using the Newton-Raphson method to the log-likelihood equations.

Each model can be tested for goodness-of-fit by, for example, the likelihood ratio Chi-squared statistic with the corresponding degrees of freedom (df). The numbers of df for the SQU, QU, and MH models are r^2-2r-1, $r(r-3)$, and $r-1$, respectively. Let $G^2(\Omega)$ denote the likelihood ratio statistic for testing the goodness-of-fit of model Ω. Thus

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

where $\hat{m}_{ij}$ is the maximum likelihood estimate of expected frequency m_{ij} under model Ω.

For the analysis of contingency tables, Lang and Agresti (1994) and Lang (1996) considered the simultaneous modeling of the joint distribution and of 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 statistic for goodness-of-fit of two models (also see Lang & Agresti, 1994; Lang, 1996; Tomizawa & Tahata, 2007; Tahata & Tomizawa, 2008).

Theorem 3

The following asymptotic equivalence holds:

$$G^2(SQU)\simeq G^2(QU)+G^2(MH). \quad (1)$$

The number of df for the SQU model equals the sum of the numbers of df for the QU and MH models.

Theorem 3 Proof

The QU model may be expressed in a log-linear form

$$\log p_{ij}=\mu^*+\alpha_i^*+\beta_j^*+ij\theta^*+\psi_{ii}^* I(i=j)$$
$$(i=1,\ldots,r;\ j=1,\ldots,r), \quad (2)$$

where $\mu^*=\log\mu$, $\alpha_i^*=\log\alpha_i$ (and so on) with $\alpha_r^*=\beta_r^*=0$ without loss of generality, and where $I(i=j)=1$ if $i=j$ and 0 otherwise. Let

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

and

$$\beta=(\mu^*,\beta_1,\beta_2,\beta_{12})',$$

where "$'$" denotes the transposed,

$$\beta_1=(\alpha_1^*,\ldots,\alpha_{r-1}^*),\ \beta_2=(\beta_1^*,\ldots,\beta_{r-1}^*),$$

and

$$\beta_{12}=(\theta^*,\psi_{11}^*,\ldots,\psi_{rr}^*).$$

The QU model is then expressed as

$$\log p = X\beta=(1_{r^2},X_1,X_2,X_{12})\beta,$$

where X is the $r^2\times 3r$ matrix and 1_s is the $s\times 1$ vector of 1 elements,

$$X_1=\begin{bmatrix} I_{r-1}\otimes 1_r \\ O_{r,r-1}\end{bmatrix};\text{the } r^2\times(r-1)\text{ matrix},$$

$$X_2=1_r\otimes\begin{bmatrix} I_{r-1} \\ 0'_{r-1}\end{bmatrix};\text{the } r^2\times(r-1)\text{ matrix},$$

and X_{12} is the $r^2\times(r+1)$ matrix, determined from (2), I_{r-1} is the $(r-1)\times(r-1)$ identity matrix, O_{st} is the $s\times t$ zero matrix, 0_s is the $s\times 1$ zero vector, and $\otimes$ denotes the Kronecker product. Note that the model matrix X is full column rank, which is $3r$. In a manner similar to Haber (1985) and Lang and Agresti (1994), the linear space spanned by the columns of the matrix X is denoted by $S(X)$ with the dimension $3r$. Let U be an $r^2 \times d_1$ full column rank matrix, where $d_1 = r^2 - 3r = r(r-3)$, such that the linear space spanned by the columns of U, that is $S(U)$, is the orthogonal complement of the space $S(X)$. Thus, $U'X = O_{d_1,3r}$. Therefore the QU model is expressed as

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

where

$$h_1(p) = U' \log p.$$

The MH model may be expressed as

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

where

$$d_2 = r - 1,\ h_2(p) = Wp,$$

and W is the $d_2 \times r^2$ matrix with

$$W = \left[I_{r-1} \otimes 1_r', O_{r-1,r}\right] - 1_r' \otimes \left[I_{r-1}, 0_{r-1}\right].$$

Therefore, $W' = X_1 - X_2$ and thus the column vectors of W' belong to the space $S(X)$, that is, $S(W') \subset S(X)$, hence, $WU = O_{d_2 d_1}$. From Theorem 2, the SQU model may be expressed as

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

where

$$d_3 = d_1 + d_2 = r^2 - 2r - 1,$$

and

$$h_3(p) = (h_1', h_2')'.$$

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 QU, MH and SQU 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, that is,

$$H_s(p) = \partial h_s(p) / \partial p'.$$

Let $\Sigma(p) = diag(p) - pp'$, where $diag(p)$ denotes a diagonal matrix with i^{th} component of p as i^{th} diagonal component. It is observed that

$$H_1(p)p = U' 1_{r^2} = 0_{d_1},$$
$$H_1(p) diag(p) = U',\ H_2(p) = W.$$

Therefore,

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

and $\Delta_3 = \Delta_1 + \Delta_2$, is obtained where

$$\Delta_s = h_s(p)'\left[H_s(p)\Sigma(p)H_s(p)'\right]^{-1} h_s(p). \tag{3}$$

From the asymptotic equivalence of the Wald statistic and the likelihood ratio statistic (Rao, 1973, Sec. 6e.3; Darroch & Silvey, 1963; Aitchison, 1962), and from (3), (1) is obtained, thus the proof is completed.

Corollary 2

The following asymptotic equivalence holds:

$$G^2(SQI) \simeq G^2(QI) + G^2(MH).$$

The number of df for the SQI model equals the sum of numbers of df for the QI and MH models.

Results

Example 1

Table 1 contains data from a case-control study investigating a possible relationship between cataracts and the use of head coverings during the summer. Each case reporting to a clinic for cataract care was matched with a control of the same gender and similar age not having a cataract. The row and column categories refer to the frequency with which the subject used head coverings.

The SQU model applied to these data has $G^2(SQU) = 10.95$ with 7 df. Thus, the SQU model fits these data well. Under this model, the maximum likelihood estimate of θ is $\hat{\theta} = 0.808$. The SQU model indicates the structure of both the S and QU models. Therefore, under the SQU model, the probability that using a head covering for one case in a pair is always or almost always, and for the control in the pair is never, is estimated to equal the probability of using a head covering for a case in the pair is never, and for control in the pair is always or almost always.

For local 2×2 tables that do not contain a cell on the main diagonal, the odds that using head covering for a case in a pair is $s+1$ instead of s is estimated to be $\hat{\theta} = 0.808$ times when that for the control in the pair is $t+1$ than when it is t. For $i < j$ and $s < t$ with $i \neq s$, $i \neq t$, $j \neq s$, $j \neq t$, the odds that the using head covering case in a pair is j instead of i is estimated to be $(0.808)^{(j-i)(t-s)}$ times higher when that for the control in the pair is t than when it is s. For example, the odds that the using a head covering for a case in a pair is never instead of frequency is estimated to be $0.426\,[= (0.808)^4]$ times higher when that for control in the pair is occasionally than when it is always or almost always.

Table 1: Case-Control Study Investigating a Possible Relationship between Cataracts and the Use of Head Coverings during the Summer

Cataract Case	Control			
	Always or Almost Always (1)	Frequently (2)	Occasionally (3)	Never (4)
Always or Almost Always (1)	29 (29.00)*	3 (4.22)	3 (5.12)	4 (6.16)
Frequently (2)	5 (4.22)	0 (0.00)	1 (1.16)	1 (1.12)
Occasionally (3)	9 (5.12)	0 (1.16)	2 (2.00)	0 (0.72)
Never (4)	7 (6.16)	3 (1.12)	1 (0.72)	0 (0.00)

*Note: The parenthesized values are the maximum likelihood estimates of expected frequencies under the SQU model.

Example 2

Table 2 contains data from the Los Angeles study of endometrial cancer. These data were obtained from 59 matched pairs using four dose levels of conjugated oestrogen: (1) none, (2) 0.1-0.299 mg, (3) 0.3-0.625 mg, and (4) 0.626+ mg.

Table 3 shows that the SQU and MH models fit the data poorly while the QU model fits the data well. From Theorem 2, the poor fit of the SQU model may be said to be caused by the influence of the lack of structure of the MH model rather than the QU model. Because the QU model fits the data in Table 2 well, under this model, the cell probabilities $\{p_{ij}\}$ have a uniform local association for cells off the main diagonal of the table.

Table 2: Average Doses of Conjugated Oestrogen Used By Cases and Matched Controls: Los Angeles Endometrial Cancer Study (Breslow & Day, 1980, p. 185)

Average Dose for Case (mg/day)	Average Dose for Control (mg/day)				Total
	0 (1)	0.1-0.299 (2)	0.3-0.625 (3)	0.625+ (4)	
0 (1)	6	2	3	1	12
0.1-0.299 (2)	9	4	2	1	16
0.3-0.625 (3)	9	2	3	1	15
0.625+ (4)	12	1	2	1	16
Total	36	9	10	4	59

Table 3: Likelihood Ratio Chi-squared Values G^2 for Models Applied to Tables 1 and 2

Applied Models	Table1		Table2	
	Degrees of Freedom	G^2	Degrees of Freedom	G^2
QI	5	6.99	5	0.77
QU	4	6.52	4	0.69
SQI	8	11.56	8	19.98^*
SQU	7	10.95	7	19.86^*
S	6	8.29	6	19.27^*
QS	3	3.85	3	0.46
MH	3	4.38	3	19.12^*

* means are significant at the 0.05 level

Conclusion

This article gives the decomposition of the SQU model and shows the orthogonality of test statistics. As observed in the examples, Theorem 2 would be useful for explaining the reason for the poor fit when the SQU model fits the data poorly.

From Theorem 3 it may be noted that the likelihood ratio statistic for testing goodness-of-fit of the SQU model - assuming that the QU model holds true - is $G^2(SQU) - G^2(QU)$ and this is asymptotically equivalent to the likelihood ratio statistic for testing goodness-of-fit of the MH model, that is, $G^2(MH)$. Namely, $G^2(SQU) - G^2(QU)$ would be used for testing goodness-of-fit of the MH model.

Suppose that model Ω_3 holds if and only if both models Ω_1 and Ω_2 hold, where the number of df for Ω_3 equals the sum of numbers of df for Ω_1 and Ω_2. Darroch and Silvey (1963) described that (i) when the asymptotic equivalence, then

$$G^2(\Omega_3) \simeq G^2(\Omega_1) + G^2(\Omega_2) \qquad (4)$$

holds, if both Ω_1 and Ω_2 are accepted (at the α significance level) with high probability, then Ω_3 would be accepted; however, (ii) when equation (4) does not hold, it is possible for an incompatible situation to arise where both Ω_1 and Ω_2 are accepted with high probability but Ω_3 is rejected with high probability (Darroch and Silvey (1963) show an interesting example). For the orthogonal decomposition of the SQU model into the QU and MH models, such an incompatible situation would not arise in terms of Theorem 3. Therefore, the orthogonal decomposition of the SQU model obtained herein is useful for analyzing data.

References

Agresti, A. (1996). *An introduction to categorical data analysis*. New York: Wiley.

Aitchison, J. (1962). Large-sample restricted parametric tests. *Journal of the Royal Statistical Society*, *Ser. B*, *24*, 234-250.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis: theory and practice.* Cambridge, MA: The MIT Press.

Bowker, A. H. (1948). A test for symmetry in contingency tables. *Journal of the American Statistical Association*, *43*, 572-574.

Breslow, N. E., & Day, N. E. (1980). *Statistical methods in cancer research, Vol. I: The analysis of case-control studies.* Lyon, France: International Agency for Research on Cancer.

Caussinus, H. (1965). Contribution à l'analyse statistique des tableaux de corré lation. *Annales de la Faculté des Sciences de l'Université de Toulouse*, *29*, 77-182.

Darroch, J. N., & Ratcliff, D. (1972). Generalized iterative scaling for log-linear models. *Annals of Mathematical Statistics*, *43*, 1470-1480.

Darroch, J. N., & Silvey, S. D. (1963). On testing more than one hypothesis. *Annals of Mathematical Statistics*, *34*, 555-567.

Goodman, L. A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association*, *74*, 537-552.

Goodman, L. A. (1985). The analysis of cross-classified data having ordered and/or unordered categories: association models, correlation models, and asymmetry models for contingency tables with or without missing entries. *Annals of Statistics*, *13*, 10-69.

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., & 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: Wiley.

Read, C. B. (1977). Partitioning chi-square in contingency tables: a teaching approach. *Communications in Statistics-Theory and Methods*, *6*, 553-562.

Stuart, A. (1955). A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika*, *42*, 412-416.

Tahata, K., & Tomizawa, S. (2008). Orthogonal decomposition of point-symmetry for multiway tables. *Advances in Statistical Analysis*, *92*, 255-269.

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 Francaise de Statistique*, *148*, 3-36.