THE ANALYSIS OF SYMMETRY AND ASYMMETRY: ORTHOGONALITY OF DECOMPOSITION OF SYMMETRY INTO QUASI-SYMMETRY AND MARGINAL SYMMETRY FOR MULTI-WAY TABLES

Sadao TOMIZAWA[1] and Kouji TAHATA[1]

ABSTRACT

For the analysis of square contingency tables, Caussinus (1965) proposed the quasi-symmetry model and gave the theorem that the symmetry model holds if and only if both the quasi-symmetry and the marginal homogeneity models hold. Bishop, Fienberg and Holland (1975, p.307) pointed out that the similar theorem holds for three-way tables. Bhapkar and Darroch (1990) gave the similar theorem for general multi-way tables. The purpose of this paper is (1) to review some topics on various symmetry models, which include the models, the decompositions of models, and the measures of departure from models, on various symmetry and asymmetry, and (2) to show that for multi-way tables, the likelihood ratio statistic for testing goodness-of-fit of the complete symmetry model is asymptotically equivalent to the sum of those for testing the quasi-symmetry model with some order and the marginal symmetry model with the corresponding order.

Keywords: Association model, Decomposition, Independence, Likelihood ratio statistic, Marginal homogeneity, Marginal symmetry, Measure, Model, Orthogonality, Quasi-symmetry, Separability, Square contingency table, Symmetry.

RÉSUMÉ

Pour l'analyse des tableaux carrés, Caussinus (1965) a proposé le modèle de quasi-symétrie et montré qu'un tableau est symétrique si et seulement s'il satisfait à la fois quasi-symétrie et égalité des distributions marginales. Bishop, Fienberg et Holland (1975, p. 307) ont noté qu'un théorème semblable valait pour les tableaux à trois dimensions, tandis que Bhapkar et Darroch l'ont donné pour des tableaux de dimension quelconque. Le but de cet article est (1) de passer en revue les questions de symétrie, les modèles eux-mêmes, leur décomposition et les mesures d'écart au modèle pour divers concepts de symétrie et asymétrie, (2) de montrer que, pour les tableaux multiples, la statistique du rapport de vraisemblance pour tester la

1. Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Noda City, Chiba, 278-8510, Japan.
E-mail: tomizawa@is.noda.tus.ac.jp
E-mail: kouji_tahata@is.noda.tus.ac.jp

symétrie est asymptotiquement équivalente à la somme des statistiques analogues testant respectivement la quasi-symétrie d'un certain ordre et l'égalité des marges pour l'ordre correspondant.

Mots-clés : Décomposition de modèle, Homogénéité des marges, Indépendance, Modèles d'association, Orthogonalité, Quasi-symétrie, Rapport de vraisemblance, Séparabilité, Symétrie, Tables de contingence carrées.

1. Introduction

Consider the square contingency tables with the same row and column classifications, which have the non-ordered, *i.e.*, nominal categories or the ordered categories. For example, consider the data in Tables 1 and 2. The classifications in Table 1 have the nominal categories and those in Table 2 have the ordered categories.

The data in Table 1 are taken directly from Upton (1978, p.119). These data refer to the voting transitions in British elections between 1966 and 1970 of a subset of the numbers of a panel study. The particular subset to which the data refer are those panel members who remained, throughout the period 1964 to 1970, in a constituency contested by the Conservative, Labour, and Liberal parties alone. The table has symmetric classifications, these being the reported votes for the three parties, together with a reported abstention. The data in Table 2, taken from Stuart (1953), are constructed from unaided distance vision of 7477 women aged 30-39 employed in Royal Ordnance factories in Britain from 1943 to 1946. In Tables 1 and 2, many observations concentrate on the main diagonal cells. Therefore for these data, the model of independence does not hold. Namely, (a) for the data in Table 1 the voting in 1970 is strongly associated with that in 1966, and (b) for the data in Table 2 a woman's right eye grade is strongly associated with her left eye grade. Instead, we are interested in (a) for the data in Table 1 whether or not the voting transitions is symmetric, and (b) for the data in Table 2 whether or not a woman's right eye grade is symmetric to her left eye grade and in how both eyes are symmetry or asymmetry.

Generally for the analysis of square table data with the same row and column classifications, various symmetry models are usually utilized. Bowker (1948) considered the symmetry model, and Caussinus (1965) considered the quasi-symmetry model. Stuart (1955) described the marginal homogeneity model. These models describe the structures of various symmetry. Other models which describe the structures of various asymmetry instead of symmetry are given; for example, the conditional symmetry model (McCullagh, 1978), the diagonals-parameter symmetry model (Goodman, 1979a), the ordinal quasi-symmetry model (Agresti, 2002a, p.429), the extended quasi-symmetry model (Tomizawa, 1984), and the cumulative diagonals-parameter symmetry model (Tomizawa, 1993a), etc.

Caussinus (1965) gave the decomposition of the symmetry model such that the symmetry model holds if and only if both the quasi-symmetry and the marginal homogeneity models hold. In the similar way to Caussinus' decomposition, Tomizawa (1984) gave the decomposition of the conditional symmetry model into the extended quasi-symmetry, the extended marginal homogeneity, and the other models. Tomizawa also gave the decompositions of some symmetry (asymmetry) models (see Section 3).

When a model (*e.g.*, the symmetry model) does not hold, we are also interested in a measure which represents the degree of departure from the model. For example, Tomizawa (1994), and Tomizawa, Seo and Yamamoto (1998) considered the measures to represent the degree of departure from the symmetry model (see Section 5). Also, Tahata, Miyamoto and Tomizawa (2004) considered the measure of departure from the Caussinus' quasi-symmetry model.

For multi-way contingency tables, the symmetry, the quasi-symmetry and the marginal symmetry models are also considered; see, for example, Bishop, Fienberg and Holland (1975, pp.299-309), Bhapkar and Darroch (1990), and Agresti (2002a, p.440). Also, Bhapkar and Darroch (1990) extended Caussinus' (1965) decomposition of the symmetry model into the multi-way tables.

Another question concerning multi-way contingency tables is the comparison between several square tables; for example, Caussinus and Thélot (1976) considered models where the tables differ only by the asymmetric aspects while the symmetric part is common for all of them; they give an example concerning migration between French regions according to the age of migrants.

We may also be interested in the studies for the $r \times r \times r$, or more generally, the r^T tables which appear as soon as a qualitative variable (with r values) is observed T (> 2) times (qualitative longitudinal data), *e.g.*, such as in the study of socio-professional status through age or across generations, or in the study of health-related data for which the response is observed for each subject at T occasions for different times. The r^T tables may also appear such as the response is measured at T occasions (not necessarily for different time) on each subject; *e.g.*, for biomedical data, each subject may be classified as having the response (with r values) by each of T drugs. For such r^T table data, we are also interested in applying, *e.g.*, generalizations of symmetry, quasi-symmetry, and marginal homogeneity. For example, see Bishop *et al.* (1975, pp.299-309), Agresti (2002a, chap.11), Tomizawa (1995a), Tomizawa and Makii (2001), Tahata, Katakura and Tomizawa (2007), and Yamamoto (2004).

The purpose of this paper is (1) to review some topics on various symmetry models (Sections 2-6), and (2) to show the orthogonality of decomposition for goodness-of-fit test of the symmetry model for multi-way tables (Section 7). Section 2 describes various symmetry models, Section 3 describes the decompositions of some symmetry models, Section 4 gives the examples, Section 5 describes the measures of some symmetry models, Section 6 describes the symmetry, the quasi-symmetry and the marginal symmetry models, and the relationship among the models, for multi-way tables, and Section 7 shows the orthogonality of test statistic of the symmetry model for multi-way tables.

2. Models

2.1. Symmetry and marginal homogeneity models

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 ($i = 1, \ldots, r; j = 1, \ldots, r$).

Bowker (1948) considered the symmetry model defined by

$$p_{ij} = \psi_{ij} \quad (i \neq j),$$

where $\psi_{ij} = \psi_{ji}$. This indicates that the probability that an observation will fall in row category i and column category j is equal to the probability that the observation falls in row category j and column category i. Namely this describes a structure of symmetry of the probabilities $\{p_{ij}\}$ with respect to the main diagonal of the table. The symmetry model is described by many statisticians; see, for example, Bishop *et al.* (1975, p.282), Caussinus (1965), Bhapkar (1979), Goodman (1985), McCullagh (1977), van der Heijden, Falguerolles and Leeuw (1989), van der Heijden and Mooijaart (1995), Agresti and Natarajan (2001), Agresti (2002a, p.424), Andersen (1994, p.320), Everitt (1992, p.142), and Tomizawa (1993a), etc. For the log-linear form of the symmetry model, see Section 6 and, *e.g.*, Bishop *et al.* (1975, p.282).

The marginal homogeneity 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). This model indicates that the row marginal distribution is identical to the column marginal distribution. For testing goodness-of-fit of this model, *e.g.*, Stuart (1955) and Bhapkar (1966) gave the Wald type test statistics, Ireland, Ku and Kullback (1969) gave the minimum discrimination information statistics, Bishop *et al.* (1975, p.294) gave the statistic based on the maximum likelihood estimates of expected frequencies, and Agresti (1983b) gave the Mann-Whitney type test statistic.

Miyamoto, Tahata, Ebie and Tomizawa (2006) considered a marginal inhomogeneity model for nominal data, defined by

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

where $|\Delta_i| = \Delta$ and $\Delta > 0$. This indicates that the odds, $p_{i\cdot}/p_{\cdot i}$ ($i = 1, \ldots, r$), are equal to e^{Δ} for some i and $e^{-\Delta}$ for the other i. Note that this model is used when the marginal homogeneity model does not hold. [See Sections 3.2 and 3.3 for some models of marginal inhomogeneity for ordinal data].

2.2. Quasi-symmetry model

Caussinus (1965) considered the quasi-symmetry model defined by

$$p_{ij} = \alpha_i \beta_j \psi_{ij} \quad (i \neq j),$$

where $\psi_{ij} = \psi_{ji}$. A special case of this model with $\{\alpha_i = \beta_i\}$ is the symmetry model. Caussinus' (1965) paper is quoted in many literatures; for example, Agresti (1983a, 1995), Agresti and Lang (1993), Bishop *et al.* (1975, p.286), Goodman (1979b), Bhapkar (1979), Bhapkar and Darroch (1990), Becker (1990), McCullagh (1982), van der Heijden *et al.* (1989), Bartolucci, Forcina and Dardanoni (2001), Haberman (1979, p.490), Plackett (1981, p.78), Clogg and Shihadeh (1994, p.66), Tomizawa (1987, 1989, 1992a, 1992b), and Tahata *et al.* (2004), etc. The related topics to Caussinus' paper (*i.e.*, the quasi-symmetry model) are also on the Web site in France,

`http://www.lsp.ups-tlse.fr/Projet_QS/index.html`.

This Web site establishes a list of many papers which have quoted Caussinus' paper. Also, in the special issue of the 'Annales de la Faculté des Sciences de Toulouse (2002)', the 11 papers related to the quasi-symmetry model are collected; they include, for example, Agresti (2002b), Dossou-Gbété and Grorud (2002), Erosheva, Fienberg and Junker (2002), Falguerolles and van der Heijden (2002), Goodman (2002), McCullagh (2002), and Caussinus (2002), etc.

Denote the odds ratio for rows i and j $(> i)$, and columns s and t $(> s)$ by $\theta_{ij;st}$. Thus $\theta_{ij;st} = (p_{is}p_{jt})/(p_{js}p_{it})$. Using odds ratios, the quasi-symmetry model is further expressed as

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

Therefore this model has characterization in terms of symmetry of odds ratios (though the symmetry model has characterization in terms of symmetry of cell probabilities). Goodman (1979b) referred to this model as the symmetric association model. We note that the symmetry of odds ratios also holds under the symmetry model, and that the independence model indicates $\theta_{ij;st} = 1$ $(i < j; s < t)$.

Let X_1 and X_2 denote the row and column variables, respectively. Also let p^c_{ij} $(i \neq j)$ denote the conditional probability of $(X_1, X_2) = (i, j)$ on condition that $(X_1, X_2) = (i, j)$ or (j, i). Namely $p^c_{ij} = p_{ij}/(p_{ij} + p_{ji})$, $i \neq j$. Then the quasi-symmetry model may be expressed as

$$p^c_{ij} = \frac{\gamma_i}{\gamma_i + \gamma_j} \quad (i \neq j).$$

So, this also relates to the Bradley-Terry model (Bradley and Terry, 1952; Agresti, 2002a, p.438); though the details are omitted. Caussinus (1965) also briefly indicates the relationship between quasi-symmetry and paired-comparison models. We note that the symmetry model indicates $p_{ij}^c = 1/2$ $(i \neq j)$ with $\gamma_1 = \ldots = \gamma_r$.

Each of the symmetry, the marginal homogeneity models, the marginal inhomogeneity model described in Section 2.1, and the quasi-symmetry model is invariant under the arbitrary same permutations of row and column categories. Thus it is suitable to apply these models for analyzing square tables with *nominal* categories (such as the data in Table 1), and also one may use these models for analyzing square tables with *ordered* categories (such as the data in Table 2) when one may not use the information about the category ordering.

Goodman (1979b) and Agresti (1983a) also considered various association models; for example, the null association, the uniform association, the linear-by-linear association, the row (column) effects, and the row and column effects association models, etc. Also see, for example, Goodman (1981a, 1981b, 1985, 1986), Chuang, Gheva and Odoroff (1985), Gilula and Haberman (1988), Clogg and Shihadeh (1994), and Agresti (2002a, p.369), etc.

For analyzing square contingency tables with the same row and column classifications, it may be useful to use the quasi association models which are defined only off the main diagonal cells. The quasi-uniform association model (Goodman, 1979b) is defined by

$$p_{ij} = \alpha_i \beta_j \theta^{ij} \quad (i \neq j).$$

A special case of this model with $\theta = 1$ is the quasi-independence (quasi null association) model. Obviously the quasi-independence and the quasi-uniform association models are special cases of Caussinus' quasi-symmetry model. When the known scores $\{u_i\}$ can be assigned to both of rows and columns, where $u_1 < \ldots < u_r$, the quasi linear-by-linear association model is defined by

$$p_{ij} = \alpha_i \beta_j \theta^{u_i u_j} \quad (i \neq j).$$

This model is also a special case of the quasi-symmetry model (*e.g.*, Agresti, 2002a, p.431). These models are used for analyzing square tables with *ordered* categories.

2.3. Asymmetry models

The symmetry, quasi-symmetry and marginal homogeneity models describe the structures of symmetry in various senses. In this section we shall introduce the models which describe the structures of asymmetry.

McCullagh (1978) considered the conditional symmetry model defined by

$$p_{ij} = \begin{cases} \delta \psi_{ij} & (i < j), \\ \psi_{ij} & (i \geqslant j), \end{cases}$$

where $\psi_{ij} = \psi_{ji}$ (also see Agresti, 2002a, p.431). A special case of this model obtained by putting $\delta = 1$ is the symmetry model. Note that the conditional symmetry model is equivalent to Read's (1977) proportional symmetry model and also to a log-linear model by Bishop *et al.* (1975, pp.285-286). This model indicates the symmetry of conditional probabilities such that

$$\mathrm{P}(X_1 = i, X_2 = j | X_1 < X_2) = \mathrm{P}(X_1 = j, X_2 = i | X_1 > X_2) \quad (i < j).$$

We note that McCullagh (1978) also considered other two multiplicative models which were referred to as the palindromic symmetry and the generalized palindromic symmetry models, including the symmetry and the conditional symmetry models as special cases.

Although the conditional symmetry model should be applied to the ordinal data (because this model is not invariant under the any same permutations of row and column categories), Tomizawa, Miyamoto and Funato (2004) considered the extended symmetry model which is applied to the nominal data as follows:

$$p_{ij} = \begin{cases} e^{\Delta_{ij}} \psi_{ij} & (i < j), \\ \psi_{ij} & (i \geqslant j), \end{cases}$$

where $|\Delta_{ij}| = \Delta$ and $\psi_{ij} = \psi_{ji}$. This indicates that the odds, p_{ij}/p_{ji} $(i < j)$, are equal to e^{Δ} for some $i < j$ and $e^{-\Delta}$ for the other $i < j$. Note that this model is different from the conditional symmetry model.

Goodman (1979a) considered the diagonals-parameter symmetry model defined by

$$p_{ij} = \begin{cases} \delta_{j-i} \psi_{ij} & (i < j), \\ \psi_{ij} & (i \geqslant j), \end{cases}$$

where $\psi_{ij} = \psi_{ji}$. A special case of this model obtained by putting $\delta_1 = \ldots = \delta_{r-1}$ $(= \delta)$ is the conditional symmetry model. Note that the diagonals-parameter symmetry model is applied to the ordinal data.

Agresti (1983c) considered the linear diagonals-parameter symmetry model defined by

$$p_{ij} = \begin{cases} \delta^{j-i} \psi_{ij} & (i < j), \\ \psi_{ij} & (i \geqslant j), \end{cases}$$

where $\psi_{ij} = \psi_{ji}$. This model is a special case of the diagonals-parameter symmetry model obtained by putting $\{\delta_{j-i} = \delta^{j-i}\}$, and also a special case of Caussinus' quasi-symmetry model. Using the known scores $u_1 < \ldots < u_r$, Agresti (2002a, p.429) also considered the ordinal quasi-symmetry model defined by

$$p_{ij} = \begin{cases} \delta^{u_j - u_i} \psi_{ij} & (i < j), \\ \psi_{ij} & (i \geqslant j), \end{cases}$$

where $\psi_{ij} = \psi_{ji}$. Obviously this model is a special case of the quasi-symmetry model. We note that the linear diagonals-parameter symmetry model and the ordinal quasi-symmetry model indicate the asymmetry of the cell probabilities but the symmetry of the odds ratios, and these models should be applied to the ordinal data because each of these models is not invariant under the any same permutations of row and column categories.

Let

$$G_{ij} = \sum_{s=1}^{i} \sum_{t=j}^{r} p_{st} = \mathrm{P}(X_1 \leqslant i, X_2 \geqslant j) \quad (i < j)$$

and

$$G_{ij} = \sum_{s=i}^{r} \sum_{t=1}^{j} p_{st} = \mathrm{P}(X_1 \geqslant i, X_2 \leqslant j) \quad (i > j).$$

Since the difference between the marginal cumulative probabilities $\mathrm{P}(X_1 \leqslant i)$ and $\mathrm{P}(X_2 \leqslant i)$ is equal to the difference between $G_{i,i+1}$ and $G_{i+1,i}$ for $i = 1, \ldots, r-1$, the asymmetry model for the cumulative probabilities $\{G_{ij}\}$, $i \neq j$ (instead of the cell probabilities), may be useful for making inferences such as that X_1 is stochastically less than X_2 or *vice versa*.

Tomizawa (1993a) pointed out that the multiplicative forms of the symmetry and the conditional symmetry models for $\{p_{ij}\}$ can also be expressed similarly as multiplicative forms for $\{G_{ij}\}$, $i \neq j$, namely as

$$G_{ij} = \Psi_{ij} \quad (i \neq j), \qquad p_{ii} = \Psi_{ii},$$

where $\Psi_{ij} = \Psi_{ji}$, and as

$$G_{ij} = \begin{cases} \delta\Psi_{ij} & (i < j), \\ \Psi_{ij} & (i > j), \end{cases} \qquad p_{ii} = \Psi_{ii},$$

where $\Psi_{ij} = \Psi_{ji}$, respectively; however, the diagonals-parameter symmetry model cannot be expressed as a (similar) multiplicative form for $\{G_{ij}\}$, $i \neq j$. So, Tomizawa (1993a) considered the cumulative diagonals-parameter symmetry model defined by

$$G_{ij} = \begin{cases} \Delta_{j-i}\Psi_{ij} & (i < j), \\ \Psi_{ij} & (i > j), \end{cases} \qquad p_{ii} = \Psi_{ii},$$

where $\Psi_{ij} = \Psi_{ji}$. This model states that the cumulative probability that an observation will fall in row category i or below and column category j $(> i)$ or above, is Δ_{j-i} times higher than the cumulative probability that the observation falls in column category i or below and row category j or above. Especially, $\Delta_1 \geqslant 1$ is equivalent to $\mathrm{P}(X_1 \leqslant i) \geqslant \mathrm{P}(X_2 \leqslant i)$ for every $i = 1, \ldots, r-1$. Therefore the parameter Δ_1 in this model would be useful for making inferences such as that X_1 is stochastically less than X_2 or *vice versa*. Note that the cumulative diagonals-parameter symmetry model is different from the Goodman's diagonals-parameter symmetry model.

Miyamoto, Ohtsuka and Tomizawa (2004) considered the cumulative linear diagonals-parameter symmetry and the cumulative quasi-symmetry models. For example, the cumulative quasi-symmetry model is defined by

$$G_{ij} = \alpha_i \beta_j \Psi_{ij} \quad (i \neq j), \qquad p_{ii} = \Psi_{ii},$$

where $\Psi_{ij} = \Psi_{ji}$. This model is different from the Caussinus' quasi-symmetry model though both models have similar multiplicative forms.

For the $r \times r$ table, define the global odds ratios as the odds ratio for the 2×2 tables corresponding to the $(r-1)(r-1)$ ways of collapsing the row and column classifications into dichotomies (*e.g.*, Agresti, 1984, p.20). Bartolucci *et al.* (2001) considered the quasi-global symmetry model, which indicates the symmetry of global odds ratios, and which is the analog of the Caussinus' quasi-symmetry model.

Goodman (1985) described various generalized independence models and generalized symmetry plus independence models; for example, the triangle non-symmetry plus independence model is defined by

$$p_{ij} = \begin{cases} \alpha_i \alpha_j \tau_1 & (i < j), \\ \alpha_i \alpha_j \tau_2 & (i > j), \end{cases}$$

which is a special case of the conditional symmetry model (also see Goodman, 1972, and Bishop *et al.*, 1975, pp.320-324).

Each of the models described in Section 2.3 (except the extended symmetry model in Tomizawa *et al.*, 2004) is not invariant under the arbitrary same permutations of row and column categories. Thus it is suitable to apply these models for analyzing square tables with *ordered* categories (such as the data in Table 2), however it is not suitable to apply these models for analyzing square tables with *nominal* categories (such as the data in Table 1).

3. Decompositions of models

Consider the $r \times r$ square contingency table with the same row and column classifications.

3.1. Decompositions of the symmetry model

Caussinus (1965) gave the decomposition of the symmetry model as follows:

THEOREM 1. — *The symmetry model holds if and only if both the quasi-symmetry and the marginal homogeneity models hold.*

For this decomposition, also see Bishop *et al.* (1975, p.287) and Agresti (2002a, p.429). From this theorem we see that assuming that the quasi-symmetry model holds true, the hypothesis that the symmetry model holds is equivalent to the hypothesis that the marginal homogeneity model holds. In addition, as shown in Section 7, the goodness-of-fit test statistic for the hypothesis that the symmetry model holds assuming that the quasi-symmetry model holds (*i.e.*, the marginal homogeneity model holds under the assumption) is asymptotically equivalent to the goodness-of-fit test statistic for the hypothesis that the marginal homogeneity model holds. So, in a sense of orthogonality (or separability), two components of symmetry (*i.e.*, quasi-symmetry and marginal homogeneity) are not related (see Section 7). For analyzing the data, this theorem also would be useful for seeing which structure of the quasi-symmetry and the marginal homogeneity is more lacking when the symmetry model fits the data poorly (see Section 4).

Using the ordinal quasi-symmetry model described in Section 2.3, Agresti (2002a, p.430) considered the decomposition of the symmetry model for the ordinal data as follows:

THEOREM 2. — *The symmetry model holds if and only if both the ordinal quasi-symmetry and the marginal homogeneity models hold.*

3.2. Decompositions of the conditional symmetry model

In the similar way to Caussinus' decomposition, Tomizawa (1984, 1985a, 1989, 1992a) considered some decompositions of the conditional symmetry model. In order to show one of these decompositions, we shall define three models below.

Tomizawa (1984, 1985a) considered the extended quasi-symmetry model defined by

$$p_{ij} = \alpha_i \beta_j \psi_{ij} \quad (i \neq j),$$

where $\psi_{ij} = \gamma \psi_{ji}$ $(i < j)$. A special case of this model obtained by putting $\gamma = 1$ is Caussinus' quasi-symmetry model. Using the odds ratios, this model may be expressed as

$$\theta_{ij;jk} = \gamma \theta_{jk;ij} \quad (i < j < k),$$

namely as

$$p_{ij} p_{jk} p_{ki} = \gamma p_{ji} p_{kj} p_{ik} \quad (i < j < k).$$

[We note that Tomizawa, Miyamoto, Yamamoto and Sugiyama (2007) considered the cumulative extended quasi-symmetry model which has the similar form for $\{G_{ij}\}$ instead of $\{p_{ij}\}$, $i \neq j$].

Tomizawa (1984, 1985a) also considered the extended marginal homogeneity model defined by

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

where δ is unspecified and

$$p_{i\cdot}^{(\delta)} = \delta p_{i\cdot}^{-} + p_{ii} + p_{i\cdot}^{+}, \quad p_{\cdot i}^{(\delta)} = p_{\cdot i}^{+} + p_{ii} + \delta p_{\cdot i}^{-},$$

$$p_{i\cdot}^{-} = \sum_{k=1}^{i-1} p_{ik}, \quad p_{i\cdot}^{+} = \sum_{k=i+1}^{r} p_{ik}, \quad p_{\cdot i}^{+} = \sum_{k=1}^{i-1} p_{ki}, \quad p_{\cdot i}^{-} = \sum_{k=i+1}^{r} p_{ki}.$$

This model indicates that the row marginal totals summed by multiplying the probabilities for cells in the lower-left (upper-right) triangle of the table by a common weight δ $(1/\delta)$ are equal to the column marginal totals summed in the same way. A special case of this model obtained by putting $\delta = 1$ is the marginal homogeneity model. Under this model, $\delta \geqslant 1$ is equivalent to $\mathrm{P}(X_1 \leqslant i) \geqslant \mathrm{P}(X_2 \leqslant i)$ for every $i = 1, \ldots, r-1$. Therefore the parameter δ in this model would be useful for making inferences such as that X_1 is stochastically less than X_2 or *vice versa*. We note that the extended marginal homogeneity model may be expressed as $G_{i,i+1} = \delta G_{i+1,i}$ for $i = 1, \ldots, r-1$; and also Tomizawa (1995b) considered the further generalized marginal homogeneity model as $G_{i,i+1} = \delta\phi^{i-1} G_{i+1,i}$ for $i = 1, \ldots, r-1$.

In order to consider the decomposition of the conditional symmetry model, we consider the balance model which indicates that the parameter γ in the extended quasi-symmetry model is equal to the parameter δ in the extended marginal homogeneity model when both models hold, *e.g.*, as follows;

$$\frac{\sum_{i=1}^{r-1} G_{i,i+1}}{\sum_{i=1}^{r-1} G_{i+1,i}} = \frac{\sum_{i<j<k} p_{ij}p_{jk}p_{ki}}{\sum_{i<j<k} p_{ji}p_{kj}p_{ik}}.$$

It may be not meaningful to apply only this model for the data, but this model would be useful to consider the decomposition of the conditional symmetry model. We obtain the following theorem:

THEOREM 3. — *The conditional symmetry model holds if and only if all the extended quasi-symmetry, the extended marginal homogeneity, and the balance models hold.*

Theorem 3 is an extension of Caussinus' decomposition (*i.e.*, Theorem 1).

3.3. Decompositions of other models

As extensions of marginal homogeneity model, Tomizawa (1987) considered two kinds of diagonal weighted marginal homogeneity models (say I and II). The model I is defined by

$$p_{i\cdot}^{-}(\phi) + p_{ii} + p_{i\cdot}^{+} = p_{\cdot i}^{+} + p_{ii} + p_{\cdot i}^{-}(\phi) \quad (i = 1, \ldots, r),$$

where

$$p_{i\cdot}^{-}(\phi) = \sum_{k=1}^{i-1} \phi^{i-k} p_{ik}, \quad p_{\cdot i}^{-}(\phi) = \sum_{k=i+1}^{r} \phi^{k-i} p_{ki}.$$

The model II is defined by

$$p_{i\cdot}^{-} + p_{ii} + p_{i\cdot}^{+}(\psi) = p_{\cdot i}^{+}(\psi) + p_{ii} + p_{\cdot i}^{-} \quad (i = 1, \ldots, r),$$

where

$$p_{i\cdot}^{+}(\psi) = \sum_{k=i+1}^{r} \psi^{k-i} p_{ik}, \quad p_{\cdot i}^{+}(\psi) = \sum_{k=1}^{i-1} \psi^{i-k} p_{ki}.$$

Special cases of these models obtained by putting $\phi = 1$ ($\psi = 1$) are the marginal homogeneity model. Under these models, $\phi \geqslant 1$ ($\psi \leqslant 1$) are equivalent to $\mathrm{P}(X_1 \leqslant i) \geqslant \mathrm{P}(X_2 \leqslant i)$ for every $i = 1, \ldots, r-1$. Therefore the parameter ϕ (ψ) in these models would be useful for making inferences such as that X_1 is stochastically less than X_2 or *vice versa*. Tomizawa (1987) gave the decomposition of the linear diagonals-parameter symmetry model for ordinal data as follows:

THEOREM 4. — *The linear diagonals-parameter symmetry model holds if and only if both the quasi-symmetry model and the diagonal weighted marginal homogeneity model I (or II) hold.*

Also Tomizawa (1998) gave the decomposition of the marginal homogeneity model for ordinal data as follows (though the details are omitted):

THEOREM 5. — *The marginal homogeneity model holds if and only if all the generalized marginal homogeneity, the marginal equi-means and the marginal equi-variances models hold.*

Note that the generalized marginal homogeneity model is defined in Section 3.2 and the marginal equi-means (equi-variances) models indicate the equality of means (variances) of X_1 and X_2.

Agresti (1984, p.205; 2002a, p.420) considered the marginal cumulative logistic model for ordinal data, which is an extension of the marginal homogeneity model, defined by

$$L_i^{X_1} = L_i^{X_2} + \Delta \quad (i = 1, \ldots, r-1),$$

where for $t = 1, 2$,

$$L_i^{X_t} = \mathrm{logit}[\mathrm{P}(X_t \leqslant i)] = \log \left[\frac{\mathrm{P}(X_t \leqslant i)}{1 - \mathrm{P}(X_t \leqslant i)} \right].$$

A special case of this model obtained by putting $\Delta = 0$ is the marginal homogeneity model. Miyamoto, Niibe and Tomizawa (2005) considered the conditional marginal cumulative logistic model which is defined only off the main diagonal, and using these logistic models and the marginal equi-means model, gave another decomposition of the marginal homogeneity model for square tables with ordered categories (though the details are omitted). Tahata *et al.* (2007) gave the similar decomposition of the marginal homogeneity model for multi-way tables with ordered categories.

3.4. Decompositions of the point-symmetry model

Wall and Lienert (1976) considered the point-symmetry model defined by

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

where $\psi_{ij} = \psi_{i^*j^*}$ and $l^* = r + 1 - l$ ($l = i, j$). This model states that the cell probabilities are point-symmetric with respect to the center point (when r is even) or the center cell (when r is odd); though the symmetry model states that they are line-symmetric with respect to the main diagonal of the table.

In the similar way to Caussinus' quasi-symmetry model, Tomizawa (1985b) considered the quasi-point-symmetry model defined by

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

where $\psi_{ij} = \psi_{i^* j^*}$. This model is an extension of the point-symmetry model.

Tomizawa (1985b) also considered the marginal point-symmetry model defined by

$$p_{i\cdot} = p_{i^*\cdot} \quad (i = 1, \dots, r)$$

and

$$p_{\cdot j} = p_{\cdot j^*} \quad (j = 1, \dots, r).$$

In the similar way to Caussinus' decomposition, Tomizawa (1985b) gave the following theorem:

Theorem 6. — *The point-symmetry model holds if and only if both the quasi-point-symmetry and the marginal point-symmetry models hold.*

Also see Tomizawa (1993b), and Tahata and Tomizawa (2006) for the decompositions of related point-symmetry models.

4. Examples

4.1. Example 1

Consider the voting transitions data in Table 1 with nominal categories. The row variable X_1 is the voting in 1966 and the column variable X_2 is the voting in 1970. Table 3 gives the values of the likelihood ratio chi-squared statistic for some models. We shall show simply the analysis based on the Caussinus' decomposition (*i.e.*, Theorem 1). [See Miyamoto *et al.* (2006) for more detailed analysis].

We see from Table 3 that the symmetry model fits these data poorly. The Caussinus' quasi-symmetry model fits these data well, however, the marginal homogeneity model fits poorly. From Caussinus' decomposition, we can see that the poor fit of the symmetry model is caused by the influence of the lack of structure of the marginal homogeneity rather than the quasi-symmetry.

Since the symmetry model does not fit these data well, however, the quasi-symmetry model fits well, it is seen that in these data there is not a structure of symmetry of cell probabilities $\{p_{ij}\}$ but there is a structure of symmetry of odds ratios $\{\theta_{ij;st}\}$.

4.2. Example 2

Consider the vision data in Table 2 with ordered categories. The row variable X_1 is the right eye grade and the column variable X_2 is the left eye grade. These data have been analyzed by many statisticians, *e.g.*, including Stuart (1953, 1955), Caussinus (1965), Bishop *et al.* (1975, p.284), Bhapkar (1966), McCullagh (1978), Goodman (1979a), Agresti (1983b, 1983c), White, Landis and Cooper (1982), Read (1977), Grizzle, Starmer and Koch (1969), Ireland *et al.* (1969), Miyamoto *et al.* (2004), Tomizawa (1985a, 1987, 1993a), and Tomizawa, Miyamoto and Yamamoto (2006), etc. Table 4 gives the values of the likelihood ratio chi-squared statistic for various models. We shall show simply the analysis based on the Caussinus' decomposition (*i.e.*, Theorem 1) and Theorem 3.

First, consider the models which indicate the structures of symmetry. We see from Table 4 that the symmetry model fits these data poorly. The Caussinus' quasi-symmetry model fits these data well, however, the marginal homogeneity model fits poorly. From Caussinus' decomposition, we can see that the poor fit of the symmetry model is caused by the influence of the lack of structure of the marginal homogeneity rather than the quasi-symmetry.

Next consider the models which indicate the structures of asymmetry. From Table 4 we see that the conditional symmetry and the linear diagonals-parameter symmetry models fit these data well. Also the diagonals-parameter symmetry model fits these data very well. According to the test (at the 0.05 level) based on the difference between the likelihood ratio chi-square values, the diagonals-parameter symmetry model may be preferable to the conditional symmetry model. From Theorem 3 we see that the reason why the conditional symmetry model is not necessarily so adequate for these data, is caused by the influence of the lack of structure of the extended quasi-symmetry rather than the extended marginal homogeneity and the balance models.

Moreover, the cumulative diagonals-parameter symmetry model fits these data very well. Under this model, the value of maximum likelihood estimate of parameter Δ_1 is 1.175 (Tomizawa, 1993a). Since this value is greater than 1, under this model the probability that the grade of the right eye is less than i ($i = 2, 3, 4$) is estimated to be greater than the probability that the grade of the left eye is less than i; namely, the left eye is estimated to be worse than the right eye. We omit here the analysis based on the other models and the other decompositions.

5. Measures

For the analysis of data, when the model does not hold, we are interested in applying the extend models, analyzing the residual, and also measuring the degree of departure from the model, etc. This section describes some measures for various symmetry and asymmetry.

5.1. Measure of departure from symmetry

For the $r \times r$ square contingency table with the nominal categories, Tomizawa (1994) and Tomizawa *et al.* (1998) considered the measures to represent the degree of departure from the symmetry model. Also, Tomizawa, Miyamoto and Hatanaka (2001) considered them for ordinal data. We shall describe the measure for nominal data below.

Assume that $\{p_{ij} + p_{ji} > 0, i \neq j\}$. Let

$$\delta = \sum\sum_{s \neq t} p_{st}, \quad p^*_{ij} = \frac{p_{ij}}{\delta}, \quad p^s_{ij} = \frac{1}{2}(p^*_{ij} + p^*_{ji}), \quad i \neq j.$$

Tomizawa *et al.* (1998) defined the measure by, for $\lambda > -1$,

$$\Phi^{(\lambda)} = \frac{\lambda(\lambda+1)}{2^\lambda - 1} I^{(\lambda)},$$

where

$$I^{(\lambda)} = \frac{1}{\lambda(\lambda+1)} \sum_{i=1}^{r} \sum_{\substack{j=1 \\ j \neq i}}^{r} p^*_{ij} \left[\left(\frac{p^*_{ij}}{p^s_{ij}} \right)^\lambda - 1 \right],$$

and the value at $\lambda = 0$ is taken to be the limit as $\lambda \to 0$. The value λ is chosen by the user. Note that $I^{(\lambda)}$ is the Cressie-Read power-divergence between $\{p^*_{ij}\}$ and $\{p^s_{ij}\}$, and in particular $I^{(0)}$ is the Kullback-Leibler information between them. For more details of the power-divergence see Cressie and Read (1984), and Read and Cressie (1988, p.15). This measure may be expressed as, for $\lambda > -1$,

$$\Phi^{(\lambda)} = \sum\sum_{i<j} (p^*_{ij} + p^*_{ji}) \left(1 - \frac{\lambda 2^\lambda}{2^\lambda - 1} H^{(\lambda)}_{ij} \right),$$

where

$$H^{(\lambda)}_{ij} = \frac{1}{\lambda}(1 - (p^c_{ij})^{\lambda+1} - (p^c_{ji})^{\lambda+1}),$$

$$p^c_{ij} = \frac{p_{ij}}{p_{ij} + p_{ji}}, \quad p^c_{ji} = \frac{p_{ji}}{p_{ij} + p_{ji}}.$$

Note that $H^{(\lambda)}_{ij}$ is Patil-Taillie diversity index for $\{p^c_{ij}, p^c_{ji}\}$, which includes the Shannon entropy when $\lambda = 0$; see Patil and Taillie (1982).

For each $\lambda > -1$, (i) $0 \leqslant \Phi^{(\lambda)} \leqslant 1$, (ii) there is a structure of symmetry in the $r \times r$ table, *i.e.*, $\{p_{ij} = p_{ji}\}$, if and only if $\Phi^{(\lambda)} = 0$, and (iii) the degree of departure from symmetry is largest (say, complete asymmetry), in the sense that $p^c_{ij} = 0$ (then $p^c_{ji} = 1$) or $p^c_{ji} = 0$ (then $p^c_{ij} = 1$) for all $i \neq j$, if and only if $\Phi^{(\lambda)} = 1$. According to the Cressie-Read power-divergence or Patil-Taillie index, $\Phi^{(\lambda)}$ represents the degree of asymmetry, and the degree increases as the value of $\Phi^{(\lambda)}$ increases.

Let $W^{(\lambda)}$ denote the power-divergence statistic for testing goodness-of-fit of the symmetry, *i.e.*,

$$W^{(\lambda)} = \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^{r} \sum_{j=1}^{r} n_{ij} \left[\left(\frac{2n_{ij}}{n_{ij}+n_{ji}} \right)^{\lambda} - 1 \right] \quad (-\infty < \lambda < +\infty),$$

where n_{ij} is the observed frequency in the (i,j)th cell of the table, and the values at $\lambda = -1$ and $\lambda = 0$ are taken to be the limits as $\lambda \to -1$ and $\lambda \to 0$, respectively. Note that $W^{(0)}$ and $W^{(1)}$ are the likelihood ratio chi-squared statistic and the Pearson's chi-squared statistic, respectively. Then we see

$$\hat{\Phi}^{(\lambda)} = \frac{\lambda(\lambda+1)}{2(2^{\lambda}-1)n^*} W^{(\lambda)} \quad (\lambda > -1),$$

where $n^* = \sum\sum_{i \neq j} n_{ij}$, and $\hat{\Phi}^{(\lambda)}$ is $\Phi^{(\lambda)}$ with p_{ij} replaced by $\hat{p}_{ij} = n_{ij}/n$, $n = \sum\sum n_{ij}$ (see Tomizawa *et al.*, 1998). The measure $\hat{\Phi}^{(\lambda)}$ is always in the range between 0 and 1, but the test statistic $W^{(\lambda)}$ depends on the off-diagonal observations n^*. Thus $\hat{\Phi}^{(\lambda)}$ would be better than $W^{(\lambda)}$ for comparing the degree of asymmetry in several tables (for more details see Tomizawa *et al.*, 1998, and Tomizawa *et al.*, 2001). We note that the measure $\hat{\Phi}^{(\lambda)}$ cannot be used for testing goodness-of-fit of the symmetry model although $W^{(\lambda)}$ is used only for testing it.

The measure $\hat{\Phi}^{(\lambda)}$ is useful for comparing the degree of departure from symmetry in several tables, and for measuring what degree the departure from symmetry is toward the complete asymmetry (see Tomizawa *et al.*, 1998). We point out that we cannot measure it by the goodness-of-fit test statistic of the symmetry model. We note that Yamamoto (2004) extended the measure $\Phi^{(\lambda)}$ into the multi-way table with nominal categories.

5.2. Measures of other symmetry and asymmetry

Tahata *et al.* (2004) considered the measure of departure from Caussinus' quasi-symmetry model and the Bradley-Terry model. Also, Tomizawa (1995a), Tomizawa and Makii (2001), and Tomizawa, Miyamoto and Ashihara (2003) considered the measures of departure from the marginal homogeneity model.

Some measures of departure from the asymmetry models are also considered. For example, see Tomizawa and Saitoh (1999a, 1999b) for the measures of departure from the conditional symmetry model, and see Tomizawa, Miyamoto and Yamane (2005) for the measure of departure from the diagonals-parameter symmetry model.

Yamamoto and Tomizawa (2007) gave the decomposition of measure for marginal homogeneity into the measure for the extended marginal homogeneity model and the measure from equality of marginal means.

5.3. Measure from correspondence analysis approach

For square contingency tables, van der Heijden *et al.* (1989), Gower (1977), and Constantine and Gower (1978) considered the singular value decomposition of a skew symmetric matrix (denoted by N) of which elements are residuals from the independence model, the quasi-independence model, the symmetry model and the quasi-symmetry model (being log-linear models), and analyzed the structure in residuals from the model with the correspondence analysis approach (also see Greenacre, 2000, and Dossou-Gbété and Grorud, 2002). Tomizawa and Murata (1992) considered the measure to represent the degree of residuals from the symmetry model by modifying the sum of squared singular values of matrix N (namely, by modifying the 'total inertia' in correspondence analysis).

5.4. Kappa measure of agreement and quasi-symmetry

Suppose that two raters separately classify each subject on an r response categories, and let p_{ij} denote the probability of rating i by the first rater and rating j by the second rater. Cohen (1960) proposed the measure 'kappa' describing degree of agreement between two raters defined as

$$\kappa = \frac{\sum p_{ii} - \sum p_{i\cdot}p_{\cdot i}}{1 - \sum p_{i\cdot}p_{\cdot i}}.$$

Tanner and Young (1985), Darroch and McCloud (1986), and Agresti (1988, 1989) proposed modeling the structure of agreement between raters, rather than describing it with a single summary measure. Tanner and Young (1985), Agresti (1988, 1989), and Agresti and Natarajan (2001) considered models having the structure of baseline association (null association, uniform association and linear-by-linear association) plus a main diagonal parameter.

For classification of subject h by rater a, let ρ_{hat} denote the probability that the rating is in category t. In a population of S subjects, if one assumes (i) that classifications are made independently in the sense that $p_{ij} = S^{-1}\sum_h \rho_{h1i}\rho_{h2j}$, and (ii) that $\{\rho_{hat}\}$ satisfies the condition of no three-factor interaction, then Darroch and McCloud (1986) showed that $\{p_{ij}\}$ satisfies the Caussinus' quasi-symmetry model. In this sense, reasonable models for agreement should be special cases of the quasi-symmetry model. As such a model, Agresti (1989) considered the model defined by

$$p_{ij} = \begin{cases} p_i p_j(1-\kappa) & (i \neq j), \\ p_i^2 + \kappa p_i(1-p_i) & (i = j), \end{cases}$$

where $p_i = p_{i\cdot} = p_{\cdot i}$ (also see Agresti, 2002a, p.453). This model has the structure of symmetry plus quasi-independence with kappa as parameter. Tomizawa (1992c) considered the model defined by

$$p_{ij} = \begin{cases} \phi_{ij} & (i \neq j), \\ (\kappa + (1-\kappa)\sum_{t=1}^{r} p_t^2)/r & (i = j), \end{cases}$$

where $\phi_{ij} = \phi_{ji}$ and $p_t = p_{t\cdot} = p_{\cdot t}$. This model has the structure of symmetry plus main diagonal equiprobability with kappa as parameter. Since both models are special cases of quasi-symmetry model, they would be reasonable models for agreement. Although the details are omitted, for the data from Bishop *et al.* (1975, p.397), Tomizawa (1992c) showed that both models for agreement fit well and described that three estimated κ (*i.e.*, κ estimated under the Agresti model, under the Tomizawa model, and under the saturated model) are quite close.

6. Symmetry, quasi-symmetry and marginal symmetry for multi-way tables

Consider the r^T contingency table ($T \geqslant 2$). 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 ith cell of the table. Also let X_k ($k = 1, \dots, T$) denote the kth variable. The complete symmetry (S^T) model is defined by

$$p_i = p_j$$

for any permutation $j = (j_1, \dots, j_T)$ of $i = (i_1, \dots, i_T)$. See, for instance, Bhapkar (1979), Bhapkar and Darroch (1990), Lovison (2000), and Agresti (2002a, p.440).

Denote the hth-order ($1 \leqslant h < T$) marginal probability $\mathrm{P}(X_{s_1} = i_1, \dots, X_{s_h} = i_h)$ by p_i^s, where $s = (s_1, \dots, s_h)$ and $i = (i_1, \dots, i_h)$ with $1 \leqslant s_1 < \dots < s_h \leqslant T$ and $i_k = 1, \dots, r$ ($k = 1, \dots, h$). The hth-order marginal symmetry (M_h^T) model is defined by

$$p_i^s = p_j^s = p_i^t \tag{6.1}$$

for any permutation $j = (j_1, \dots, j_h)$ of $i = (i_1, \dots, i_h)$ and for any $s = (s_1, \dots, s_h)$ and $t = (t_1, \dots, t_h)$ (Bhapkar and Darroch, 1990; Agresti, 2002a, p.440). For instance, when $T = 3$, the M_1^3 model is defined by

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

where '$\cdot$' denotes the sum; thus $p_{i\cdot\cdot} = \sum_s \sum_t p_{ist}$, etc., and the M_2^3 model is defined by

$$p_{ij\cdot} = p_{ji\cdot} = p_{i\cdot j} = p_{\cdot ij} \quad (i = 1, \dots, r; j = 1, \dots, r). \tag{6.2}$$

The S^T model may be expressed as in a log-linear form,

$$\log p_i = \lambda(i),$$

where $\lambda(i) = \lambda(j)$ for any permutation $j = (j_1, \dots, j_T)$ of $i = (i_1, \dots, i_T)$. In a three-way table, for instance, $\log p_{122} = \log p_{212} = \log p_{221} = \lambda(122)$ (Agresti, 2002a, p.440).

Bhapkar and Darroch (1990) defined the hth-order ($1 \leqslant h < T$) quasi-symmetry (Q_h^T) model, which may be expressed as

$$\log p_i = \lambda + \sum_{k=1}^{T} \lambda_k(i_k) + \sum\sum_{1 \leqslant k_1 < k_2 \leqslant T} \lambda_{k_1 k_2}(i_{k_1}, i_{k_2}) + \ldots + \sum \cdots \sum_{1 \leqslant k_1 < \ldots < k_h \leqslant T} \lambda_{k_1 \ldots k_h}(i_{k_1}, \ldots, i_{k_h}) + \lambda(i), \quad (6.3)$$

where $\lambda(i) = \lambda(j)$ for any permutation $j = (j_1, \ldots, j_T)$ of $i = (i_1, \ldots, i_T)$. For instance, when $T = 3$, the Q_1^3 model is expressed as

$$\log p_{ijk} = \lambda + \lambda_1(i) + \lambda_2(j) + \lambda_3(k) + \lambda(ijk),$$

where $\lambda(ijk) = \lambda(ikj) = \lambda(jik) = \lambda(kij) = \lambda(jki) = \lambda(kji)$; and the Q_2^3 model is expressed as

$$\log p_{ijk} = \lambda + \lambda_1(i) + \lambda_2(j) + \lambda_3(k) + \lambda_{12}(ij) + \lambda_{13}(ik) + \lambda_{23}(jk) + \lambda(ijk), \quad (6.4)$$

where $\lambda(ijk) = \ldots = \lambda(kji)$. Also see Bishop *et al.* (1975, p.303).

Bhapkar and Darroch (1990) gave the extension of Theorem 1 into multi-way tables as follows:

THEOREM 7. — *For r^T table and $1 \leqslant h < T$, the S^T model holds if and only if both the Q_h^T and M_h^T models hold.*

When $T = 2$, this theorem is identical to Theorem 1. When $T = 3$ with $h = 1$ and 2, this theorem is identical to the relation pointed out by Bishop *et al.* (1975, p.307) in which however the proof was not given.

7. Orthogonality of decomposition of test statistic for multi-way tables

Consider the r^T contingency table ($T \geqslant 2$). 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 statistics for goodness-of-fit of two models (also see Lang and Agresti, 1994; Lang, 1996; Tomizawa, 1992a, 1993b). We are now interested in whether or not, for r^T table, the test statistic for the S^T model is asymptotically equivalent to the sum of the test statistic for the Q_h^T model and that for the M_h^T model.

This section shows that for the multi-way tables, the test statistic for the S^T model is asymptotically equivalent to the sum of those for the Q_h^T and M_h^T models. We note that the Q_1^2 model is the quasi-symmetry model.

7.1. The case of $r \times r$ table

We shall consider the case of $T = 2$, *i.e.*, the $r \times r$ square table. The Q_1^2 model is expressed as

$$\log p_{ij} = \lambda + \lambda_1(i) + \lambda_2(j) + \lambda(ij) \quad (i = 1, \ldots, r; j = 1, \ldots, r), \tag{7.1}$$

where $\lambda(ij) = \lambda(ji)$. Without loss of generality, for example, we may set $\lambda_1(r) = \lambda_2(r) = \lambda(rj) = \lambda(ir) = 0$. Let

$$p = (p_{11}, \ldots, p_{1r}, p_{21}, \ldots, p_{2r}, \ldots, p_{r1}, \ldots, p_{rr})',$$
$$\beta = (\lambda, \beta_1, \beta_2, \beta_{12})',$$

where

$$\beta_1 = (\lambda_1(1), \ldots, \lambda_1(r-1)),\ \beta_2 = (\lambda_2(1), \ldots, \lambda_2(r-1)),$$

and

$$\beta_{12} = (\lambda(11), \ldots, \lambda(1, r-1), \lambda(22), \ldots, \lambda(2, r-1), \ldots, \lambda(r-1, r-1))$$

is the $1 \times r(r-1)/2$ vector of $\lambda(ij)$ for $1 \leqslant i \leqslant j \leqslant r-1$. Then the Q_1^2 model is expressed as

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

where X is the $r^2 \times K$ matrix with $K = (r^2 + 3r - 2)/2$ 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 $r^2 \times r(r-1)/2$ matrix of 1 or 0 elements, determined from (7.1), 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 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 $d_1 = r^2 - K = (r-1)(r-2)/2$, 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, $U'X = O_{d_1,K}$. Therefore the Q_1^2 model is expressed as

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

where

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

The M_1^2 model, which is the marginal homogeneity 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

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

Therefore we see that $W' = X_1 - X_2$. Thus the column vectors of W' belong to the space $S(X)$, *i.e.*, $S(W') \subset S(X)$. Hence $WU = O_{d_2 d_1}$. From Theorem 1, the S^2 model, which is the symmetry model, may be expressed as

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

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

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

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 degrees of freedoms for testing goodness-of-fit of the Q_1^2, M_1^2 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.*, $H_s(p) = \partial h_s(p)/\partial p'$. Let $\Sigma(p) = diag(p) - pp'$, where $diag(p)$ denotes a diagonal matrix with ith component of p as ith diagonal component. We see that

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

Therefore we obtain

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

Thus we obtain the following lemma.

Lemma 1. — $\Delta_3 = \Delta_1 + \Delta_2$ *holds, where*

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

Assume that a multinomial distribution applies to the $r \times r$ table. For a model, say, Ω, let $G^2(\Omega)$ denote the likelihood ratio statistic for testing goodness-of-fit of model Ω. From the asymptotic equivalence of the Wald statistic and the likelihood ratio statistic (Rao, 1973, Sec. 6e. 3) and from Lemma 1, Darroch and Silvey (1963), and Aitchison (1962), we obtain the following theorem.

Theorem 8. — *For the* $r \times r$ *table, the following asymptotic equivalence holds:*

$$G^2(S^2) \simeq G^2(Q_1^2) + G^2(M_1^2).$$

7.2. The case of $r \times r \times r$ table

We shall consider the case of $T = 3$, *i.e.*, $r \times r \times r$ table, with order $h = 2$. The Q_2^3 model is expressed as (6.4). Without loss of generality, we set $\lambda_m(r) = 0$, $\lambda_{st}(rj) = \lambda_{st}(ir) = 0$, and $\lambda(rjk) = \lambda(irk) = \lambda(ijr) = 0$ for $m = 1, 2, 3$, and $1 \leqslant s < t \leqslant 3$. Let

$$p = (p_{111}, \dots, p_{1r1}, \dots, p_{r11}, \dots, p_{rr1}, p_{112}, \dots, p_{1r2}, \dots, p_{r12}, \dots, p_{rr2}, \\ \dots, p_{11r}, \dots, p_{1rr}, \dots, p_{r1r}, \dots, p_{rrr})',$$

$$\beta = (\lambda, \beta_1, \beta_2, \beta_3, \beta_{12}, \beta_{13}, \beta_{23}, \beta_{123})',$$

where

$$\beta_m = (\lambda_m(1), \dots, \lambda_m(r-1)), \quad m = 1, 2, 3,$$

$$\beta_{st} = (\lambda_{st}(11), \dots, \lambda_{st}(1, r-1), \lambda_{st}(21), \dots, \lambda_{st}(2, r-1), \\ \dots, \lambda_{st}(r-1, 1), \dots, \lambda_{st}(r-1, r-1)), \quad 1 \leqslant s < t \leqslant 3;$$

and

$$\beta_{123} = (\lambda(111), \dots, \lambda(11, r-1), \lambda(122), \dots, \lambda(12, r-1), \dots, \lambda(222), \\ \dots, \lambda(22, r-1), \dots, \lambda(r-2, r-1, r-1), \lambda(r-1, r-1, r-1))$$

is the $1 \times (r^3 - r)/6$ vector of $\lambda(ijk)$ for $1 \leqslant i \leqslant j \leqslant k \leqslant r-1$. Then Q_2^3 model is expressed as

$$\log p = X\beta = (1_{r^3}, X_1, X_2, X_3, X_{12}, X_{13}, X_{23}, X_{123})\beta, \tag{7.2}$$

where X is the $r^3 \times K$ vector with $K = (r^3 + 18r^2 - 19r + 6)/6$,

$$X_1 = 1_r \otimes \begin{bmatrix} I_{r-1} \otimes 1_r \\ O_{r,r-1} \end{bmatrix}; \text{ the } r^3 \times (r-1) \text{ matrix},$$

$$X_2 = 1_{r^2} \otimes \begin{bmatrix} I_{r-1} \\ 0'_{r-1} \end{bmatrix}; \text{ the } r^3 \times (r-1) \text{ matrix},$$

$$X_3 = \begin{bmatrix} I_{r-1} \otimes 1_{r^2} \\ O_{r^2,r-1} \end{bmatrix}; \text{ the } r^3 \times (r-1) \text{ matrix},$$

$$X_{12} = 1_r \otimes \begin{bmatrix} I_{r-1} \otimes \begin{bmatrix} I_{r-1} \\ 0'_{r-1} \end{bmatrix} \\ O_{r,(r-1)^2} \end{bmatrix}; \text{ the } r^3 \times (r-1)^2 \text{ matrix},$$

$$X_{13} = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_{r-1} \\ O_{r^2,(r-1)^2} \end{bmatrix}; \text{ the } r^3 \times (r-1)^2 \text{ matrix},$$

where

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

with

$$A_1 = [1_r, O_{r,r-2}],\ A_{r-1} = [O_{r,r-2}, 1_r],$$

and

$$X_{23} = \begin{bmatrix} D_{11} & \cdots & D_{1,r-1} \\ & \cdots & \\ D_{r-1,1} & \cdots & D_{r-1,r-1} \\ & O_{r^2,(r-1)^2} & \end{bmatrix}; \text{ the } r^3 \times (r-1)^2 \text{ matrix,}$$

where

$$D_{ij} = 1_r \otimes E_{ji}; \text{ the } r^2 \times (r-1) \text{ matrix,}$$

E_{ji} is the $r \times (r-1)$ matrix with a 1 in the (j,i)th element and 0's elsewhere, and X_{123} is the $r^3 \times (r^3 - r)/6$ matrix of 1 or 0 elements, determined from (6.4). Let U be an $r^3 \times d_1$, where $d_1 = r^3 - K = (r-1)(r-2)(5r-3)/6$, 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, $U'X = O_{d_1,K}$. Therefore the Q_2^3 model is expressed as

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

where

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

The M_2^3 model, which is defined by (6.2), may be expressed as

$$p_{i\cdot\cdot} = p_{\cdot i\cdot} = p_{\cdot\cdot i} \quad (i = 1, \ldots, r-1), \tag{7.3}$$
$$p_{ij\cdot} = p_{i\cdot j} \quad (i = 1, \ldots, r-1; j = 1, \ldots, r-1), \tag{7.4}$$
$$p_{ij\cdot} = p_{\cdot ij} \quad (i = 1, \ldots, r-1; j = 1, \ldots, r-1), \tag{7.5}$$
$$p_{ij\cdot} = p_{ji\cdot} \quad (i = 1, \ldots, r-1; j = 1, \ldots, r-1). \tag{7.6}$$

The equation (7.3), which is the M_1^3 model, is expressed as

$$W_1 p = 0_{2(r-1)},$$

where W_1 is the $2(r-1) \times r^3$ matrix with

$$W_1' = [X_1,\ X_1] - [X_2,\ X_3].$$

The equations (7.4) and (7.5) are expressed as

$$W_2 p = 0_{2(r-1)^2},$$

where W_2 is the $2(r-1)^2 \times r^3$ matrix with

$$W_2' = [X_{12},\ X_{12}] - [X_{13},\ X_{23}]\,.$$

The equation (7.6) is expressed as

$$W_3 p = 0_{(r-1)(r-2)/2},$$

where W_3 is the $(r-1)(r-2)/2 \times r^3$ matrix with

$$W_3' = A_1 - A_2,$$

$$\begin{aligned} A_1 &= [a_{12}, a_{13}, \ldots, a_{1,r-1}, a_{23}, \ldots, a_{2,r-1}, \ldots, a_{r-2,r-1}], \\ A_2 &= [a_{21}, a_{31}, \ldots, a_{r-1,1}, a_{32}, \ldots, a_{r-1,2}, \ldots, a_{r-1,r-2}], \end{aligned}$$

and a_{ij} is the $r^3 \times 1$ vector, being one of column vectors in X_{12} for (7.2), shouldering $\lambda_{12}(ij)$. Thus, the M_2^3 model is expressed as

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

where $d_2 = (r-1)(5r-2)/2$,

$$h_2(p) = Wp,\ W = \begin{bmatrix} W_1 \\ W_2 \\ W_3 \end{bmatrix}.$$

All column vectors of W' belong to the space $S(X)$, *i.e.*, $S(W') \subset S(X)$. Therefore, in the similar manner to the case of $r \times r$ table (in Section 7.1), we see

$$H_1(p)\Sigma(p)H_2(p)' = O_{d_1 d_2}.$$

Although the detail is omitted, the similar result is also obtained for order $h = 1$, *i.e.*, for the Q_1^3 and M_1^3 models. Therefore we obtain

THEOREM 9. — *For the $r \times r \times r$ table, the following asymptotic equivalence holds: for order $h = 1, 2$,*

$$G^2(S^3) \simeq G^2(Q_h^3) + G^2(M_h^3).$$

7.3. The case of r^T table

Consider the r^T table. For order h with $1 \leqslant h < T$, the Q_h^T model is expressed as (6.3). Without loss of generality, we may set $\lambda_{k_1 \ldots k_l}(i_{k_1}, \ldots, i_{k_l}) = 0$ if $i_t = r$ for some i_t of $(i_{k_1}, \ldots, i_{k_l})$, where $1 \leqslant l \leqslant h$, and $\lambda(i_1, \ldots, i_T) = 0$ if $i_{m_1} = \ldots = i_{m_v} = r$ for some $i_{m_1}, \ldots, i_{m_v}$ of $(i_1, \ldots, i_T)$, where $T-h \leqslant v \leqslant T$. In the similar way to the cases of $r \times r$ and $r \times r \times r$ tables, for $1 \leqslant h < T$, the Q_h^T model is expressed as the form,

$$\log p = X\beta. \tag{7.7}$$

From (6.1), the M_1^T model is expressed as

$$p_{(i)}^1 = p_{(i)}^2 = \ldots = p_{(i)}^T \quad (i = 1, \ldots, r-1), \tag{7.8}$$

where $p_{(i)}^s = \mathrm{P}(X_s = i)$. The M_2^T model may be expressed as the M_1^T model, *i.e.*, (7.8), plus

$$p_{(i,j)}^{(s_1,s_2)} = p_{(j,i)}^{(s_1,s_2)} = p_{(i,j)}^{(t_1,t_2)}$$

for $i = 1, \ldots, r-1; j = 1, \ldots, r-1; 1 \leqslant s_1 < s_2 \leqslant T$ and $1 \leqslant t_1 < t_2 \leqslant T$. Similarly, the M_h^T model may be expressed as the M_{h-1}^T model plus

$$p_i^s = p_j^s = p_i^t$$

for any permutation $j = (j_1, \ldots, j_h)$ of $i = (i_1, \ldots, i_h)$, where $i_k = 1, \ldots, r-1$ $(k = 1, \ldots, h)$, and for $s = (s_1, \ldots, s_h)$ and $t = (t_1, \ldots, t_h)$, where $1 \leqslant s_1 < \ldots < s_h \leqslant T$ and $1 \leqslant t_1 < \ldots < t_h \leqslant T$. Then the M_h^T model is expressed as

$$Wp = 0_{d_2}, \tag{7.9}$$

where

$$d_2 = \sum_{u=0}^{h} \binom{T}{u} (r-1)^u - \sum_{u=0}^{h} \binom{r-2+u}{u}, \tag{7.10}$$

though the detail is omitted. In (7.9), for instance, consider one of the restrictions for the M_h^T model for the mth $(1 \leqslant m \leqslant h)$ marginal probabilities such that

$$p_i^s - p_i^t = 0$$

for $i = (i_1, \ldots, i_m)$, where $i_k = 1, \ldots, r-1$ $(k = 1, \ldots, m)$, and for $s = (s_1, \ldots, s_m)$ and $t = (t_1, \ldots, t_m)$, where $1 \leqslant s_1 < \ldots < s_m \leqslant T$ and $1 \leqslant t_1 < \ldots < t_m \leqslant T$. This restriction is expressed as

$$w_l p = 0,$$

where w_l is a $1 \times r^T$ vector being one of the row vectors of W in (7.9). Let $a_{s_1 \ldots s_m}(i_1, \ldots, i_m)$ and $a_{t_1 \ldots t_m}(i_1, \ldots, i_m)$ be the $r^T \times 1$ vectors, being column vectors in X for (7.7), shouldering $\lambda_{s_1\ldots s_m}(i_1,\ldots,i_m)$ and $\lambda_{t_1\ldots t_m}(i_1,\ldots,i_m)$ in β, respectively. Then we see

$$w'_l = a_{s_1\ldots s_m}(i_1,\ldots,i_m) - a_{t_1\ldots t_m}(i_1,\ldots,i_m).$$

Thus the vector w'_l belongs to the space $S(X)$. Similarly, all column vectors of W' belong to the space $S(X)$, *i.e.*, $S(W') \subset S(X)$. Therefore, in the similar manner to the cases of $r \times r$ and $r \times r \times r$ tables, we obtain the following theorem.

Theorem 10. — *For the r^T table, the following asymptotic equivalence holds: for order h $(1 \leqslant h < T)$,*

$$G^2(S^T) \simeq G^2(Q_h^T) + G^2(M_h^T).$$

The numbers of degrees of freedoms for testing goodness-of-fit of the Q_h^T, M_h^T and S^T models are d_1, d_2 and d_3, respectively, where d_2 is given by (7.10),

$$d_1 = r^T - \sum_{u=0}^{h} \binom{T}{u} (r-1)^u - \sum_{u=h+1}^{T} \binom{r-2+u}{u}$$

and

$$d_3 = r^T - \binom{r-1+T}{T},$$

though the detail is omitted. Note that $d_3 = d_1 + d_2$ since it is easily seen that

$$\sum_{u=0}^{T} \binom{r-2+u}{u} = \binom{r-1+T}{T}.$$

7.4. Concluding remarks

We point out from Theorem 10 that for instance, the likelihood ratio statistic for testing goodness-of-fit of the S^T model assuming that the Q_h^T model holds true is $G^2(S^T) - G^2(Q_h^T)$ and this is asymptotically equivalent to the likelihood ratio statistic for testing goodness-of-fit of the M_h^T model, *i.e.*, $G^2(M_h^T)$.

We see that for each of the data in Tables 1 and 2 the value of $G^2(S^2)$ is very close to the sum of the values of $G^2(Q_1^2)$ and $G^2(M_1^2)$ (see Tables 3 and 4). As described in Section 4 we see by Theorem 8 that for each of the data in Tables 1 and 2 the poor fit of the symmetry model is caused by the influence of the lack of structure of the marginal homogeneity rather than the quasi-symmetry.

The decomposition of test statistic is interesting not only from the testing point of view but also sheds some light on the "decomposition" of symmetry into quasi-symmetry and marginal homogeneity.

Acknowledgements

The authors would like to express their sincere thanks to Professor Henri Caussinus, an associate editor and the referee for their many helpful suggestions.

References

Agresti A. (1983a). A survey of strategies for modeling cross-classifications having ordinal variables. *Journal of the American Statistical Association*, **78**, 184-198.

Agresti A. (1983b). Testing marginal homogeneity for ordinal categorical variables. *Biometrics*, **39**, 505-510.

Agresti A. (1983c). A simple diagonals-parameter symmetry and quasi-symmetry model. *Statistics and Probability Letters*, **1**, 313-316.

Agresti A. (1984). *Analysis of Ordinal Categorical Data*. Wiley, New York.

Agresti A. (1988). A model for agreement between ratings on an ordinal scale. *Biometrics*, **44**, 539-548.

Agresti A. (1989). An agreement model with kappa as parameter. *Statistics and Probability Letters*, **7**, 271-273.

Agresti A. (1995). Logit models and related quasi-symmetric log-linear models for comparing responses to similar items in a survey. *Sociological Methods and Research*, **24**, 68-95.

Agresti A. (2002a). *Categorical Data Analysis*, 2nd edition. Wiley, New York.

Agresti A. (2002b). Links between binary and multi-category logit item response models and quasi-symmetric loglinear models. *Annales de la Faculté des Sciences de Toulouse*, **11**, 443-454.

Agresti A. and Lang J. B. (1993). Quasi-symmetric latent class models, with application to rater agreement. *Biometrics*, **49**, 131-139.

Agresti A. and Natarajan R. (2001). Modeling clustered ordered categorical data: A survey. *International Statistical Review*, **69**, 345-371.

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

Andersen E. B. (1994). *The Statistical Analysis of Categorical Data*, 3rd edition. Springer, Berlin.

Bartolucci F., Forcina A. and Dardanoni V. (2001). Positive quadrant dependence and marginal modeling in two-way tables with ordered margins. *Journal of the American Statistical Association*, **96**, 1497-1505.

Becker M. P. (1990). Quasisymmetric models for the analysis of square contingency tables. *Journal of the Royal Statistical Society, Series B*, **52**, 369-378.

Bhapkar V. P. (1966). A note on the equivalence of two test criteria for hypotheses in categorical data. *Journal of the American Statistical Association*, **61**, 228-235.

Bhapkar V. P. (1979). On tests of marginal symmetry and quasi-symmetry in two and three-dimensional contingency tables. *Biometrics*, **35**, 417-426.

Bhapkar V. P. and Darroch J. N. (1990). Marginal symmetry and quasi symmetry of general order. *Journal of Multivariate Analysis*, **34**, 173-184.

Bishop Y. M. M., Fienberg S. E. and Holland P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice.* The MIT Press, Cambridge, Massachusetts.

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

BRADLEY R. A. and TERRY M. E. (1952). Rank analysis of incomplete block designs I. The method of paired comparisons. *Biometrika*, **39**, 324-345.

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

CAUSSINUS H. (2002). Some concluding observations. *Annales de la Faculté des Sciences de Toulouse*, **11**, 587-591.

CAUSSINUS H. and THÉLOT C. (1976). Note complémentaire sur l'analyse statistique des migrations. *Annales de l'INSEE*, **22-23**, 135-146.

CHUANG C., GHEVA D. and ODOROFF C. (1985). Methods for diagnosing multiplicative-interaction models for two-way contingency tables. *Communications in Statistics-Theory and Methods*, **14**, 2057-2080.

CLOGG C. C. and SHIHADEH E. S. (1994). *Statistical Models for Ordinal Variables.* Sage Publications, California.

COHEN J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**, 37-46.

CONSTANTINE A. G. and GOWER J. C. (1978). Graphical representation of asymmetric matrices. *Applied Statistics*, **27**, 297-304.

CRESSIE N. A. C. and READ T. R. C. (1984). Multinomial goodness-of-fit tests. *Journal of the Royal Statistical Society, Series B*, **46**, 440-464.

DARROCH J. N. and MCCLOUD P. I. (1986). Category distinguishability and observer agreement. *Australian Journal of Statistics*, **28**, 371-388.

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

DOSSOU-GBÉTÉ S. and GRORUD A. (2002). Biplots for matched two-way tables. *Annales de la Faculté des Sciences de Toulouse*, **11**, 469-483.

EROSHEVA E. A., FIENBERG S. E. and JUNKER B. W. (2002). Alternative statistical models and representations for large sparse multi-dimensional contingency tables. *Annales de la Faculté des Sciences de Toulouse*, **11**, 485-505.

EVERITT B. S. (1992). *The Analysis of Contingency Tables*, 2nd edition. Chapman and Hall, London.

DE FALGUEROLLES A. and van der HEIJDEN, P. G. M. (2002). Reduced rank quasi-symmetry and quasi-skew symmetry: a generalized bi-linear model approach. *Annales de la Faculté des Sciences de Toulouse*, **11**, 507-524.

GILULA Z. and HABERMAN S. J. (1988). The analysis of multivariate contingency tables by restricted canonical and restricted association models. *Journal of the American Statistical Association*, **83**, 760-771.

GOODMAN L. A. (1972). Some multiplicative models for the analysis of cross-classified data. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 649-696.

GOODMAN L. A. (1979a). Multiplicative models for square contingency tables with ordered categories. *Biometrika*, **66**, 413-418.

GOODMAN L. A. (1979b). 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. (1981a). Association models and canonical correlation in the analysis of cross-classifications having ordered categories. *Journal of the American Statistical Association*, **76**, 320-334.

GOODMAN L. A. (1981b). Association models and the bivariate normal for contingency tables with ordered categories. *Biometrika*, **68**, 347-355.

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.

GOODMAN L. A. (1986). Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables. *International Statistical Review*, **54**, 243-309.

GOODMAN L. A. (2002). Contributions to the statistical analysis of contingency tables: Notes on quasi-symmetry, quasi-independence, log-linear models, log-bilinear models, and correspondence analysis models. *Annales de la Faculté des Sciences de Toulouse*, **11**, 525-540.

GOWER J. C. (1977). The analysis of asymmetry and orthogonality. *In Recent Developments in Statistics, J. R. Barra et al. ed., North-Holland, Amsterdam*, 109-123.

GREENACRE M. (2000). Correspondence analysis of square asymmetric matrices. *Applied Statistics*, **49**, 297-310.

GRIZZLE J. E., STARMER C. F. and KOCH G. G. (1969). Analysis of categorical data by linear models. *Biometrics*, **25**, 489-504.

HABER M. (1985). Maximum likelihood methods for linear and log-linear models in categorical data. *Computational Statistics and Data Analysis*, **3**, 1-10.

HABERMAN S. J. (1979). *Analysis of Qualitative Data, Volume 2.* Academic Press, New York.

IRELAND C. T., KU H. H. and KULLBACK S. (1969). Symmetry and marginal homogeneity of an $r \times r$ contingency table. *Journal of the American Statistical Association*, **64**, 1323-1341.

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.

LOVISON G. (2000). Generalized symmetry models for hypercubic concordance tables. *International Statistical Review*, **68**, 323-338.

MCCULLAGH P. (1977). A logistic model for paired comparisons with ordered categories. *Biometrika*, **64**, 449-453.

MCCULLAGH P. (1978). A class of parametric models for the analysis of square contingency tables with ordered categories. *Biometrika*, **65**, 413-418.

MCCULLAGH P. (1982). Some applications of quasisymmetry. *Biometrika*, **69**, 303-308.

MCCULLAGH P. (2002). Quasi-symmetry and representation theory. *Annales de la Faculté des Sciences de Toulouse*, **11**, 541-561.

MIYAMOTO N., NIIBE K. and TOMIZAWA S. (2005). Decompositions of marginal homogeneity model using cumulative logistic models for square contingency tables with ordered categories. *Austrian Journal of Statistics*, **34**, 361-373.

MIYAMOTO N., OHTSUKA W. and TOMIZAWA S. (2004). Linear diagonals-parameter symmetry and quasi-symmetry models for cumulative probabilities in square contingency tables with ordered categories. *Biometrical Journal*, **46**, 664-674.

Miyamoto N., Tahata K., Ebie H. and Tomizawa S. (2006). Marginal inhomogeneity models for square contingency tables with nominal categories. *Journal of Applied Statistics*, **33**, 203-215.

Patil G. P. and Taillie C. (1982). Diversity as a concept and its measurement. *Journal of the American Statistical Association*, **77**, 548-561.

Plackett P. L. (1981). *The Analysis of Categorical Data*, 2nd edition. Charles Griffin, London.

Rao C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd edition. Wiley, New York.

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

Read T. R. C. and Cressie N. A. C. (1988). *Goodness-of-Fit Statistics for Discrete Multivariate Data*. Springer, New York.

Stuart A. (1953). The estimation and comparison of strengths of association in contingency tables. *Biometrika*, **40**, 105-110.

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

Tahata K., Katakura S. and Tomizawa S. (2007). Decompositions of marginal homogeneity model using cumulative logistic models for multi-way contingency tables. *Revstat*, **5**, 163-176.

Tahata K., Miyamoto N. and Tomizawa S. (2004). Measure of departure from quasi-symmetry and Bradley-Terry models for square contingency tables with nominal categories. *Journal of the Korean Statistical Society*, **33**, 129-147.

Tahata K. and Tomizawa S. (2006). Decompositions for extended double symmetry model in square contingency tables with ordered categories. *Journal of the Japan Statistical Society*, **36**, 91-106.

Tanner M. A. and Young M. A. (1985). Modeling agreement among raters. *Journal of the American Statistical Association*, **80**, 175-180.

Tomizawa S. (1984). Three kinds of decompositions for the conditional symmetry model in a square contingency table. *Journal of the Japan Statistical Society*, **14**, 35-42.

Tomizawa S. (1985a). Analysis of data in square contingency tables with ordered categories using the conditional symmetry model and its decomposed models. *Environmental Health Perspectives*, **63**, 235-239.

Tomizawa S. (1985b). The decompositions for point-symmetry models in two-way contingency tables. *Biometrical Journal*, **27**, 895-905.

Tomizawa S. (1987). Diagonal weighted marginal homogeneity models and decompositions for linear diagonals-parameter symmetry model. *Communications in Statistics-Theory and Methods*, **16**, 477-488.

Tomizawa S. (1989). Decompositions for conditional symmetry model into palindromic symmetry and modified marginal homogeneity models. *Australian Journal of Statistics*, **31**, 287-296.

Tomizawa S. (1992a). A decomposition of conditional symmetry model and separability of its test statistic for square contingency tables. *Sankhyā, Series B*, **54**, 36-41.

Tomizawa S. (1992b). Multiplicative models with further restrictions on the usual symmetry model. *Communications in Statistics-Theory and Methods*, **21**, 693-710.

TOMIZAWA S. (1992c). An agreement model having structure of symmetry plus main-diagonal equiprobability. *Journal of the Korean Statistical Society*, **21**, 179-185.

TOMIZAWA S. (1993a). Diagonals-parameter symmetry model for cumulative probabilities in square contingency tables with ordered categories. *Biometrics*, **49**, 883-887.

TOMIZAWA S. (1993b). Orthogonal decomposition of point-symmetry model for two-way contingency tables. *Journal of Statistical Planning and Inference*, **36**, 91-100.

TOMIZAWA S. (1994). Two kinds of measures of departure from symmetry in square contingency tables having nominal categories. *Statistica Sinica*, **4**, 325-334.

TOMIZAWA S. (1995a). Measures of departure from marginal homogeneity for contingency tables with nominal categories. *The Statistician*, **44**, 425-439.

TOMIZAWA S. (1995b). A generalization of the marginal homogeneity model for square contingency tables with ordered categories. *Journal of Educational and Behavioral Statistics*, **20**, 349-360.

TOMIZAWA S. (1998). A decomposition of the marginal homogeneity model into three models for square contingency tables with ordered categories. *Sankhyā, Series B*, **60**, 293-300.

TOMIZAWA S. and MAKII T. (2001). Generalized measures of departure from marginal homogeneity for contingency tables with nominal categories. *Journal of Statistical Research*, **35**, 1-24.

TOMIZAWA S., MIYAMOTO N. and ASHIHARA N. (2003). Measure of departure from marginal homogeneity for square contingency tables having ordered categories. *Behaviormetrika*, **30**, 173-193.

TOMIZAWA S., MIYAMOTO N. and FUNATO R. (2004). Conditional difference asymmetry model for square contingency tables with nominal categories. *Journal of Applied Statistics*, **31**, 271-277.

TOMIZAWA S., MIYAMOTO N. and HATANAKA Y. (2001). Measure of asymmetry for square contingency tables having ordered categories. *Australian and New Zealand Journal of Statistics*, **43**, 335-349.

TOMIZAWA S., MIYAMOTO N., YAMAMOTO K. and SUGIYAMA A. (2007). Extensions of linear diagonals-parameter symmetry and quasi-symmetry models for cumulative probabilities in square contingency tables. *Statistica Neerlandica*, **61**, 273-283.

TOMIZAWA S., MIYAMOTO N. and YAMANE S. (2005). Power-divergence-type measure of departure from diagonals-parameter symmetry for square contingency tables with ordered categories. *Statistics*, **39**, 107-115.

TOMIZAWA S., MIYAMOTO N. and YAMAMOTO K. (2006). Decomposition for polynomial cumulative symmetry model in square contingency tables with ordered categories. *Metron*, **64**, 303-314.

TOMIZAWA S. and MURATA M. (1992). Gauss discrepancy type measure of degree of residuals from symmetry for square contingency tables. *Journal of the Korean Statistical Society*, **21**, 59-69.

TOMIZAWA S. and SAITOH K. (1999a). Measure of departure from conditional symmetry for square contingency tables with ordered categories. *Journal of the Japan Statistical Society*, **29**, 65-78.

TOMIZAWA S. and SAITOH K. (1999b). Kullback-Leibler information type measure of departure from conditional symmetry and decomposition of measure from symmetry for contingency tables. *Calcutta Statistical Association Bulletin*, **49**, 31-39.

Tomizawa S., Seo T. and Yamamoto H. (1998). Power-divergence-type measure of departure from symmetry for square contingency tables that have nominal categories. *Journal of Applied Statistics*, **25**, 387-398.

Upton G. J. G. (1978). *The Analysis of Cross-tabulated Data*. Wiley, New York.

van der Heijden P. G. M., de Falguerolles A. and de Leeuw J. (1989). A combined approach to contingency table analysis using correspondence analysis and log-linear analysis. *Applied Statistics*, **38**, 249-292.

van der Heijden P. G. M. and Mooijaart A. (1995). Some new log-bilinear models for the analysis of asymmetry in a square contingency table. *Sociological Methods and Research*, **24**, 7-29.

Wall K. D. and Lienert G. A. (1976). A test for point-symmetry in J-dimensional contingency-cubes. *Biometrical Journal*, **18**, 259-264.

White A. A., Landis J. R. and Cooper M. M. (1982). A note on the equivalence of several marginal homogeneity test criteria for categorical data. *International Statistical Review*, **50**, 27-34.

Yamamoto H. (2004). A measure of departure from symmetry for multi-way contingency tables with nominal categories. *Japanese Journal of Biometrics*, **25**, 69-88.

Yamamoto K. and Tomizawa S. (2007). Decomposition of measure for marginal homogeneity in square contingency tables with ordered categories. *Austrian Journal of Statistics*, **36**, 105-114.

Annex

TABLE 1. — Voting changes between 1966 and 1970 British Elections; from Upton (1978, p.119).

	1970				
1966	Conservative	Labour	Liberal	Abstention	Total
Conservative	68	1	1	7	77
Labour	12	60	5	10	87
Liberal	12	3	13	2	30
Abstention	8	2	3	6	19
Total	100	66	22	25	213

TABLE 2. — Unaided distance vision of 7477 women aged 30-39 employed in Royal Ordnance factories in Britain from 1943 to 1946; from Stuart (1953).

Right eye	Left eye grade				
grade	Best (1)	Second (2)	Third (3)	Worst (4)	Total
Best (1)	1520	266	124	66	1976
Second (2)	234	1512	432	78	2256
Third (3)	117	362	1772	205	2456
Worst (4)	36	82	179	492	789
Total	1907	2222	2507	841	7477

TABLE 3. — Likelihood ratio chi-square values of models applied to the data in Table 1.

Applied models	Degrees of freedom	Likelihood ratio chi-square
Symmetry	6	28.54*
Quasi-symmetry	3	4.20
Marginal homogeneity	3	24.97*

* means significant at the 0.05 level.

TABLE 4. — Likelihood ratio chi-square values of models applied to the data in Table 2.

Applied models	Degrees of freedom	Likelihood ratio chi-square
Symmetry	6	19.25*
Quasi-symmetry	3	7.27
Marginal homogeneity	3	11.99*
Conditional symmetry	5	7.35
Linear diagonals-parameter symmetry	5	7.28
Diagonals-parameter symmetry	3	0.50
Cumulative diagonals-parameter symmetry	3	0.02
Extended quasi-symmetry	2	6.82*
Extended marginal homogeneity	2	0.005
Diagonal weighted marginal homogeneity I	2	0.005
Diagonal weighted marginal homogeneity II	2	0.015
Balance	1	0.11

* means significant at the 0.05 level.