Advances in Quasi-Symmetry for Square Contingency Tables

Kouji Tahata [ID]

Department of Information Sciences, Faculty of Science and Technology, Tokyo University of Science, Yamazaki, Noda-shi 278-8510, Chiba, Japan; kouji_tahata@rs.tus.ac.jp

Abstract: Contingency tables highlight relationships between categorical variables. Typically, the symmetry or marginal homogeneity of a square contingency table is evaluated. The original symmetry model often does not accurately fit a dataset due to its restrictions. Caussinus proposed a quasi-symmetry model which served as a bridge between symmetry and marginal homogeneity in square contingency tables. This study significantly influenced methodological developments in the statistical analysis of categorical data. Herein recent advances in quasi-symmetry are reviewed with an emphasis on four topics related to the author's results: (1) modeling based on the f-divergence, (2) the necessary and sufficient condition of symmetry, (3) partition of test statistics for symmetry, and (4) measure of the departure from symmetry. The asymmetry model based on f-divergence enables us to express various asymmetries. Additionally, these models are useful to derive the necessary and sufficient conditions of symmetry with desirable properties. This review may be useful to consider the statistical modeling and the measure of symmetry for contingency tables with the same classifications.

Keywords: decomposition; f-divergence; marginal homogeneity; measure; symmetry

1. Introduction

Contingency tables play important roles in various fields, as they highlight relationships between categorical variables. Typically, the analysis of contingency tables is interested in whether row and column variables are independent. If the independence hypothesis is rejected, then the association between the variables is of interest. Various coefficients have been proposed to measure the association, such as *gamma*, *Yule's Q*, *Kendall's tau-b*, *Kendall's tau*, and *Somers' d*. See (Bishop et al. [1], Ch. 11) and (Agresti [2], pp. 184–192). Bishop et al. [1], Agresti [2], and Kateri [3] present overviews of contingency table analysis. Additionally, Kateri [4] reviewed the ϕ-divergence association models related to the independence model for two-way contingency tables. Fujisawa and Tahata [5] proposed the generalized asymmetry plus quasi-uniform association model. These works in the literature provide reviews of contingency table analysis based on association models.

Contingency tables with the same row and column classifications are often called *square* contingency table. Square contingency table data arise many fields. For example, unaided distance vision data, social mobility data, father–son matched educational level data, longitudinal data in biomedical research, and so on. Typically, the analysis of square contingency tables considers the issue of symmetry rather than independence (null association) because observations tend to concentrate on or near the main diagonal. Therefore, many statisticians have considered various symmetry and asymmetry models, and modeling of symmetry is one of the important topics for the analysis of square contingency tables. For example, Bowker [6] proposed a test for the hypothesis of symmetry (S). The S model indicates the symmetry structure of cell probabilities. Additionally, Stuart [7] provided a large-sample test for marginal homogeneity (MH). The MH model indicates the equivalence of marginal probabilities.

Caussinus [8] developed quasi-symmetry (QS). The QS model bridges S and MH in square tables. The QS model indicates the symmetry of the odds ratio with respect to the

main diagonal of the square contingency table. If the S model holds, then the MH model holds. However, the inverse is not always true. The QS model can be used to show that the S model holds if and only if both the QS and MH models hold. This finding has influenced the methodological developments in statistical analysis of categorical data. See Section 3 for more details.

A special issue of *Annales de la Faculté des Sciences de Toulouse, Mathématiques* was published in 2002. It contained papers written by internationally distinguished authors on topics related to QS. Agresti [9] described some generalizations of the QS model that have similar connections with generalizations of the Rasch model. Goodman [10] commented relationship between the QS model and the quasi-independence model. McCullagh [11] set out to list all hereditary sub-representations by real-valued square matrices, and to explain how these may be used in model construction. There are Fienberg and van der Heijden [12], Bergsma and Rudas [13], Dossou-Gbété and Grorud [14], Erosheva et al. [15], De Falguerolles and van der Heijden [16], Stigler [17], Thélot [18], and Caussinus [19] in addition to those in the special issue.

Since then, numerous studies have treated QS from a variety of perspectives. For square contingency tables, Tomizawa [20] introduced modeling based on the cumulative probabilities. Miyamoto et al. [21] proposed the cumulative QS model for square contingency tables with ordinal categories. Kateri et al. [22] considered generalized QS models. To evaluate the goodness of fit, Booth et al. [23] adapted a network algorithm to test QS in square tables. Krampe et al. [24] proposed a method based on algebraic statistics combined with Markov chain Monte Carlo (MCMC) methods for the ordinal QS (OQS) model. Additionally, various topics related to the QS model are discussed in mathematical statistics. For example, see Rapallo [25], Pardo and Martin [26], and Gottard et al. [27]. Tomizawa and Tahata [28] reviewed some topics on various symmetry models and showed the property of test statistics of symmetry for multi-way contingency tables. Tahata and Tomizawa [29] reviewed various models of symmetry and asymmetry and presented the relationships among models.

The topics related to the QS model have been discussed in several papers for the last 5 years. Bocci and Rapallo [30] reviewed why the synergy between algebraic statistics and quasi-independence has been fruitful. Additionally, see Khan and Tewari [31]. Altun [32] considered various symmetry and asymmetry models for square contingency tables with ordinal categories. Tahata et al. [33] proposed the model selection via the penalized likelihood approach. The symmetry models for square contingency tables are applied to the cross classified a single nucleotide polymorphism (SNP) interactions data in Karadağ et al. [34]. Altunay and Yilmaz [35] proposed two novel log-linear models to measure the degree of accumulation of the neutral option over the contingency tables based on Likert-type items. Additionally, Ando [36,37,38] proposed models that indicate the structure of asymmetry and gave the decompositions of the models. For square contingency tables with ordinal categories, the models which have various ordered scores were compared in Ando [39].

This review focuses on the further developments of modeling and properties of symmetry in recent years. Additionally, some comments which are not described in referenced papers are added. Herein four topics related to Tahata [40] and Tahata et al. [41] are reviewed: (1) modeling based on the f-divergence, (2) the necessary and sufficient condition of symmetry, (3) partition of test statistics for symmetry, and (4) the measure of departure from symmetry.

The rest of this paper is organized as follows. Section 2 demonstrates modeling based on the f-divergence. Section 3 reviews the necessary and sufficient condition of symmetry. The result given by Caussinus [8] is included as a special case. Aitchison [42], Darroch and Silvey [43], Read [44], Lang and Agresti [45], and Lang [46] discussed the partitioning of goodness of fit statistics. Section 4 presents an overview of the main results presented in Tomizawa and Tahata [28]. Caussinus' QS model has a good property from a partitioning point of view. Section 5 introduces the measure of departure from symmetry. Section 6

discusses the relationships between the measure and asymmetry model. Finally, Section 7 contains concluding remarks.

2. Modeling Based on the f-Divergence

Consider an $R \times R$ square contingency table with the same row and column classifications. Let X and Y denote the row and column variables, respectively, and let π_{ij} denote the probability that an observation will fall in the cell (i, j) of the table.

The S model, which is defined as $\pi_{ij} = \pi_{ji}$ for $i = 1, \dots, R; j = 1, \dots, R$, indicates the symmetry structure of cell probabilities (Bowker [6]). Additionally, the MH model indicates the equivalence of marginal probabilities and is defined as $\pi_{i\bullet} = \pi_{\bullet i}$ for $i = 1, \dots, R$ where $\pi_{i\bullet} = \sum_j \pi_{ij}$ and $\pi_{\bullet j} = \sum_i \pi_{ij}$ (Stuart [7]). The QS model is defined as

$$\pi_{ij}\pi_{jk}\pi_{ki} = \pi_{ji}\pi_{kj}\pi_{ik} \quad (1 \le i < j < k \le R). \tag{1}$$

From (1), the QS model can be expressed using the odds ratio as

$$\theta_{(i<j;j<k)} = \theta_{(j<k;i<j)} \quad (1 \le i < j < k \le R),$$

where $\theta_{(i<j;j<k)} = (\pi_{ij}\pi_{jk})/(\pi_{jj}\pi_{ik})$ and $\theta_{(j<k;i<j)} = (\pi_{ji}\pi_{kj})/(\pi_{ki}\pi_{jj})$.

This indicates the symmetry of the odds ratio with respect to the main diagonal of the square contingency table (Caussinus [8]). For the analysis of square contingency tables, symmetry models have been proposed by using various ideas, for example, McCullagh [47], Agresti [48], and Tomizawa [49].

Ireland et al. [50], (Bishop et al. [1], pp. 345–346) and Gilula et al. [51] presented the method for model generation. In particular, Kateri and Papaioannou [52] applied the method for the modeling of symmetry for square contingency tables and proposed a family of models, which includes the QS model as a special case. This family is derived by minimizing the f-divergence under certain conditions. Additionally, Tahata [40] extended the results to the modeling of the asymmetry structure of cell probabilities. This section reviews asymmetry models based on the f-divergence.

Let $\boldsymbol{\pi} = (\pi_{ij})$ and $\boldsymbol{q} = (q_{ij})$ be two discrete finite bivariate probability distributions. (Csiszár and Shields [53], Section 4) introduced the f-divergence between $\boldsymbol{\pi}$ and $\boldsymbol{q}$. Let f be a convex function on $(0, +\infty)$ with $f(1) = 0$. The f-divergence of a distribution $\boldsymbol{\pi}$ from $\boldsymbol{q}$ is defined as

$$I^C(\boldsymbol{\pi} : \boldsymbol{q}) = \sum_i \sum_j q_{ij} f\left(\frac{\pi_{ij}}{q_{ij}}\right).$$

Here, we take $f(0) = \lim_{t\to 0} f(t)$, $0 \cdot f(0/0) = 0$, and $0 \cdot f(a/0) = a \lim_{t\to\infty}[f(t)/t]$.

Let f be a twice differentiable and strictly convex function, and $\pi_{ij}^S = (\pi_{ij} + \pi_{ji})/2$. Kateri and Papaioannou [52] proposed the generalized QS (denoted by QS[f]) model. It is defined as

$$\pi_{ij} = \pi_{ij}^S F^{-1}(a_i + \gamma_{ij}) \quad (i = 1, \dots, R; j = 1, \dots, R), \tag{2}$$

where $\gamma_{ij} = \gamma_{ji}$ and $F(t) = f'(t)$. The QS[f] model is the closest to the S model in terms of the f-divergence under the condition where the marginals $\pi_{i\bullet}$ (or $\pi_{\bullet i}$) for $i = 1, \dots, R$ and the sums $\pi_{ij} + \pi_{ji} = 2\pi_{ij}^S$ for $i = 1, \dots, R; j = 1, \dots, R$ are given. Kateri and Papaioannou [52] noted that if $f(t) = t \log(t)$, $t > 0$, then the f-divergence is reduced to the Kullback–Leibler divergence (Kullback and Leibler [54]), and the QS[f] model is equivalent to the QS model. Namely, the QS model is the closest to the S model in terms of the Kullback–Leibler divergence under this condition. Additionally, see Kateri [4], Kateri and Agresti [55], Kateri [56], and Tahata [40].

Next, consider an $R \times R$ square contingency table with ordered categories. Let $\{u_i\}$ be a set of known scores $u_1 \le u_2 \le \cdots \le u_R$ (with $u_1 < u_R$). Replacing $\{a_i\}$ by $\{\alpha u_i\}$ in

the QS[f] model is the ordinal QS[f] (OQS[f]) model (Kateri and Agresti [55]). The OQS[f] model is the closest to the S model in terms of the f-divergence under the condition where $\sum_i u_i \pi_{i\bullet}$ (or $\sum_i u_i \pi_{\bullet i}$) and the sums $\{\pi_{ij} + \pi_{ji}\}$ are given. If $f(t) = t\log(t)$, $t > 0$, then (1) the OQS[f] model with $\{u_i = i\}$ is equivalent to the linear diagonals parameter S (LDPS) model (Agresti [48]) and (2) the OQS[f] model is equivalent to the OQS model (Agresti [9], pp. 236–238).

Tahata [40] proposed a model to fill the gap between the QS[f] and OQS[f] models. It should be noted that $\{u_i\}$ is a set of known scores $u_1 < u_2 < \cdots < u_R$. For a given k ($k = 1, \ldots, R-1$) the asymmetry model based on the f-divergence (AS$_k$[f]) is defined as

$$\pi_{ij} = \pi_{ij}^S F^{-1}\left(\sum_{h=1}^{k} u_i^h \alpha_h + \gamma_{ij}\right) \quad (i = 1, \ldots, R; j = 1, \ldots, R),$$

where $\gamma_{ij} = \gamma_{ji}$ and $F(t) = f'(t)$. From the relation $\pi_{ij}^S = (\pi_{ij} + \pi_{ji})/2$, the parameters of the AS$_k$[f] model must satisfy

$$F^{-1}\left(\sum_{h=1}^{k} u_i^h \alpha_h + \gamma_{ij}\right) + F^{-1}\left(\sum_{h=1}^{k} u_j^h \alpha_h + \gamma_{ji}\right) = 2.$$

The AS$_k$[f] model is the closest to the S model in terms of the f-divergence under the condition where $\sum_i u_i^h \pi_{i\bullet}$ (or $\sum_i u_i^h \pi_{\bullet i}$) for $h = 1, \ldots, k$ and the sums $\pi_{ij} + \pi_{ji}$ for $i = 1, \ldots, R; j = 1, \ldots, R$ are given. Additionally, the AS$_k$[f] model can be expressed as

$$\pi_{ij} = \pi_{ij}^S F^{-1}\left(\sum_{h=0}^{k} u_i^h \alpha_h + \gamma_{ij}\right) \quad (i = 1, \ldots, R; j = 1, \ldots, R), \tag{3}$$

where $\gamma_{ij} = \gamma_{ji}$. For example, $\alpha_0 = 1$ can be set without a loss of generality. There is a one-to-one transformation between a set of $\{a_1, a_2, \ldots, a_R\}$ in Equation (2) and that of $\{\alpha_0, \alpha_1, \ldots, \alpha_{R-1}\}$ in Equation (3) when $k = R - 1$. When $\alpha_1 = \cdots = \alpha_k = 0$, the AS$_k$[$f$] model is reduced to the S model. It should be noted that (1) the AS$_1$[f] model is reduced to the OQS[f] model, (2) the AS$_k$[f] model ($k = 2, 3, \ldots, R - 2$) is an extension of the OQS[f] model, and (3) the AS$_{R-1}$[f] model is reduced to the QS[f] model.

If $f(t) = t\log(t)$ for $t > 0$, then the AS$_k$[f] model becomes

$$\pi_{ij} = \pi_{ij}^S \exp\left[\sum_{h=1}^{k} u_i^h \alpha_h + \gamma_{ij} - 1\right] \quad (i = 1, \ldots, R; j = 1, \ldots, R), \tag{4}$$

where $\gamma_{ij} = \gamma_{ji}$. Then under Equation (4)

$$\frac{\pi_{ij}}{\pi_{ji}} = \prod_{h=1}^{k} \beta_h^{u_i^h - u_j^h} \quad (i < j), \tag{5}$$

where $\beta_h = \exp[\alpha_h]$. Equation (5) with $\{u_i = i\}$ is the kth linear asymmetry (LS$_k$) model proposed by Tahata and Tomizawa [57]. Therefore, the LS$_k$ model is the closest to the S model in terms of the Kullback–Leibler divergence under the condition where $\sum_i i^h \pi_{i\bullet}$ (or $\sum_i i^h \pi_{\bullet i}$) for $h = 1, \ldots, k$ and the sums $\{\pi_{ij} + \pi_{ji}\}$ are given.

Let the conditional probability that an observation will fall in cell (i, j) when the observation falls in cell (i, j) or (j, i) be denoted by π_{ij}^c. That is, $\pi_{ij}^c = \pi_{ij}/(\pi_{ij} + \pi_{ji})$. When $f(t) = (1 - t)^2$, the AS$_k$[f] model becomes

$$\pi_{ij} = \pi_{ij}^S\left(\frac{\sum_{h=1}^{k} u_i^h \alpha_h + \gamma_{ij}}{2} + 1\right) \quad (i = 1, \ldots, R; j = 1, \ldots, R), \tag{6}$$

where $\gamma_{ij} = \gamma_{ji}$. Then under Equation (6)

$$\pi_{ij}^c - \pi_{ji}^c = \sum_{h=1}^{k} \left(u_i^h - u_j^h \right) \beta_h \quad (i < j),$$ (7)

where $\beta_h = \alpha_h/4$ with $\sum_h (u_j^h - u_i^h)\beta_h < 1$. This model is the closest to the S model when the divergence is measured by the Pearson distance.

Equation (5) can be expressed as

$$\log \frac{\pi_{ij}^c}{\pi_{ji}^c} = \sum_{h=1}^{k} \left(u_i^h - u_j^h \right) \alpha_h \quad (i < j).$$

Namely, the $AS_k[f]$ model with $f(t) = t \log(t)$ indicates that the log odds is expressed as a polynomial of (α_h). On the other hand, the $AS_k[f]$ model with $f = (1 - t)^2$ (namely, Equation (6)) indicates that the difference between symmetric conditional probabilities is expressed as a polynomial of (β_h) from Equation (7). Let $g(x) = 2x$ and the composition of functions F and g be denoted by H. That is, $H(x) = F(2x)$. Then the $AS_k[f]$ model can be expressed as

$$H(\pi_{ij}^c) = \sum_{h=1}^{k} u_i^h \alpha_h + \gamma_{ij} \quad (i = 1, \ldots, R; j = 1, \ldots, R),$$

where $\gamma_{ij} = \gamma_{ji}$. This formula indicates that a H-transformation of conditional probability (π_{ij}^c) has a linear combination of parameters. For $i < j$,

$$H(\pi_{ij}^c) - H(\pi_{ji}^c) = \sum_{h=1}^{k} \left(u_i^h - u_j^h \right) \alpha_h.$$

Therefore, the $AS_k[f]$ model is characterized as the structure with symmetric conditional probabilities (π_{ij}^c, π_{ji}^c). The structure depends on the function f. The function f is chosen by the user. Namely, we can apply various asymmetry models to the given dataset by changing the function f. Namely, modeling based on f-divergence enable to construct various asymmetries by using the function f.

Fujisawa and Tahata [58] proposed the extended $AS_k[f]$ model for square contingency tables with ordinal categories. Additionally, Yoshimoto et al. [59] discussed the quasi point-symmetry models based on the f-divergence. For two-way contingency tables, the ϕ-divergence (f-divergence) association models are a family of models which includes the association and correlation models as special cases. Kateri [4] presented this family of models and demonstrated the role of ϕ-divergence in building this family. Additionally, Kateri [56] developed the new families of ϕ-divergence generalized QS models.

3. Necessary and Sufficient Condition of Symmetry

The S model rarely fits the given dataset due to its strong restrictions. More relaxed models, such as the QS and MH models, are often applied to a dataset. Here, we are interested in an extension of the S model, which indicates the asymmetry structure of cell probabilities. Caussinus [8] noted that the S model holds if and only if both the QS and MH models hold. Caussinus' result is useful to deduce the reason for a poor fit of the S model, and Bishop et al. [1] (Section 8.2.3) derived its proof. Let H_M denote model M. Then

$$H_S = H_{QS} \cap H_{MH}.$$ (8)

In this section, the necessary and sufficient conditions for the S model are reviewed.

The conditional S (CS) model (Read [44] and McCullagh [47]) is defined as $\pi_{ij} = \gamma\pi_{ji}$ for $i < j$. The CS model with $\gamma = 1$ is reduced to the S model. See also (Bishop et al. [1], pp. 285–286).

Consider the global symmetry (GS) model, which is defined as

$$\sum_{i<j}\sum \pi_{ij} = \sum_{i>j}\sum \pi_{ij}.$$

Namely, this model indicates that $\Pr(X < Y)$ is equal to $\Pr(X > Y)$. If both the CS and GS models hold, then the S model holds. Since the converse holds, Read [44] highlighted that

$$H_S = H_{CS} \cap H_{GS}. \tag{9}$$

The LDPS model fits well when there is an underlying bivariate normal distribution (Agresti [48]). Additionally, Tomizawa [49] considered an extended LDPS (ELDPS) model and reported the relationship with bivariate normal distribution. Tahata and Tomizawa [57] considered the LS_k model, which is defined by Equation (5) with $\{u_i = i\}$. This model is a generalization of the LDPS and ELDPS models. That is, the LDPS and ELDPS models are the LS_1 and LS_2 models, respectively. It should be noted that

$$H_S \subset H_{LS_1} (= H_{LDPS}) \subset H_{LS_2} (= H_{ELDPS}) \subset \cdots \subset H_{LS_{R-2}} \subset H_{LS_{R-1}} (= H_{QS}). \tag{10}$$

For $m = 1, \ldots, R - 2$, the difference in the degrees of freedom (df) between H_{LS_m} and $H_{LS_{m+1}}$ is one.

On the other hand, for a given positive integer k ($k = 1, \ldots, R - 1$), the kth moment equality (ME$_k$) model is defined as

$$E(X^h) = E(Y^h) \quad (h = 1, \ldots, k),$$

where $E(X^h) = \sum_i i^h \pi_{i\bullet}$ and $E(Y^h) = \sum_i i^h \pi_{\bullet i}$. For example, the ME$_1$ model indicates that the mean of X is equal to that of Y, and the ME$_2$ model indicates the equality of the mean and variance for X and Y. Tahata and Tomizawa [60] proved that the MH model is equivalent to the ME$_{R-1}$ model. Therefore,

$$H_S \subset H_{ME_{R-1}} (= H_{MH}) \subset H_{ME_{R-2}} \subset \cdots \subset H_{ME_2} \subset H_{ME_1}. \tag{11}$$

For $m = 1, \ldots, R - 2$, the difference in df between H_{ME_m} and $H_{ME_{m+1}}$ is one. From the Caussinus' result, Equations (10) and (11) provide various relations. For example

$$H_S = H_{LS_k} \cap H_{ME_{R-1}} (= H_{MH}) \quad (k = 1, \ldots, R - 2),$$

and

$$H_S = H_{LS_1} (= H_{LDPS}) \cap H_{ME_k} \quad (k = 2, \ldots, R - 1).$$

(Agresti [2], p. 261) and Kateri and Agresti [55] reported similar comments. However, it should be noted that $H_{LS_k} \cap H_{MH}$ and $H_{LDPS} \cap H_{ME_k}$ have more redundant conditions than the S model.

Let (π_{ij}^*) denote the cell probabilities satisfying both the LS_k and ME$_k$ models. The LS_k model is expressed as

$$\log \pi_{ij}^* = \sum_{h=1}^{k} \left(i^h \log u_{1(h)} + j^h \log u_{2(h)} \right) + \log \psi_{ij} \quad (i = 1, \ldots, R; j = 1, \ldots, R),$$

where $\psi_{ij} = \psi_{ji}$. Additionally, the ME_k model is expressed as

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

where $\mu_1^{*(h)} = \sum_i i^h \pi_{i\bullet}^*$ and $\mu_2^{*(h)} = \sum_i i^h \pi_{\bullet i}^*$. By considering I^C with $f(t) = t \log(t)$ (namely, the Kullback–Leibler divergence) to measure the difference between (π_{ij}^*) and (π_{ji}^*),

$$
\begin{aligned}
\sum_i \sum_j \pi_{ij}^* \log \frac{\pi_{ij}^*}{\pi_{ji}^*} &= \sum_i \sum_j \pi_{ij}^* \left\{ \sum_{h=1}^k \left(i^h - j^h \right) \log u_{1(h)} + \left(j^h - i^h \right) \log u_{2(h)} \right\} \\
&= \sum_{h=1}^k \left\{ \left(\mu_1^{*(h)} - \mu_2^{*(h)} \right) \log u_{1(h)} + \left(\mu_2^{*(h)} - \mu_1^{*(h)} \right) \log u_{2(h)} \right\} \\
&= 0.
\end{aligned}
$$

That is, $\pi_{ij}^* = \pi_{ji}^*$ holds for $i = 1, \ldots, R; j = 1, \ldots, R$ from the property of the Kullback–Leibler divergence. Additionally, if the S model holds, then both LS_k and ME_k models hold clearly. Therefore, for a given positive integer k, these lead to

$$H_S = H_{LS_k} \cap H_{ME_k}. \tag{12}$$

Tahata [40] extended the result using the $AS_k[f]$ model $(k = 1, \ldots, R - 1)$. Consider the generalized ME_k (denoted by GME_k) model defined as

$$\sum_i u_i^h \pi_{i\bullet} = \sum_i u_i^h \pi_{\bullet i} \quad (h = 1, \ldots, k),$$

with a set of known scores $u_1 < u_2 < \cdots < u_R$. Tahata [40] proved the following result:

$$H_S = H_{AS_k[f]} \cap H_{GME_k}. \tag{13}$$

Tahata et al. [61] proposed the extended LS_k (ELS_k) model, which is defined as

$$\frac{\pi_{ij}}{\pi_{ji}} = \gamma \prod_{h=1}^k \theta_h^{j^h - i^h} \quad (i < j).$$

The ELS_k model with $\gamma = 1$ is reduced to the LS_k model. It should be noted that the ELS_{R-1} model is equal to the extended QS (EQS) model (Tomizawa [62]), which indicates the asymmetry structure of the odds ratio. The QS model indicates the symmetry structure of the odds ratios. The OQS (LDPS) model indicates both the QS model and the asymmetry structure of cell probabilities and is a special case of the QS model. The EQS model is an extension of the QS model. Tahata et al. [61] noted that

$$H_S = H_{ELS_k} \cap H_{ME_k} \cap H_{GS}. \tag{14}$$

Fujisawa and Tahata [58] proposed the model, which is a generalization of the ELS_k model using the f-divergence and gave the generalization of Equation (14).

4. Partition of Test Statistics for Symmetry

Consider the situation where the analyst has found hypothesis H_{M_3} unacceptable and has turned their attention to an examination of components H_{M_1} and H_{M_2} such that $H_{M_3} = H_{M_1} \cap H_{M_2}$. Let $T(M)$ denote the statistic for testing the goodness of fit of model M. $T(M)$ is asymptotically distributed as a chi-square distribution with the corresponding df. Aitchison [42] noted that the possibility of partitioning the test statistic for H_{M_3} into components for testing H_{M_1} and H_{M_2} must be investigated. Namely, $T(M_3) = T(M_1) + T(M_2)$. When hypotheses H_{M_1} and H_{M_2} are *separable*, partitioning is

possible. *Separable* implies that the restricted tests of H_{M_3} against H_{M_2} and that of H_{M_3} against H_{M_1} use the same critical regions as the unrestricted test of H_{M_1} and that of H_{M_2}, respectively. It should be noted that df for H_{M_3} (denoted by df_3) is equal to the sum of df for H_{M_1} and H_{M_2}. That is, $df_3 = df_1 + df_2$. The acceptance of H_{M_1} and H_{M_2} separately means that $T(M_1) < c_{df_1}(\alpha)$ and $T(M_2) < c_{df_2}(\alpha)$ where α is the significance level and c denotes the corresponding critical value. Then $T(M_1) + T(M_2) < c_{df_1}(\alpha) + c_{df_2}(\alpha)$. Under the separable hypotheses, hypothesis H_{M_3} is rejected when $T(M_1) + T(M_2) > c_{df_3}(\alpha)$. The acceptance of H_{M_1} and H_{M_2} usually results in the acceptance of H_{M_3} because $c_{df_1}(\alpha) + c_{df_2}(\alpha)$ rarely exceeds $c_{df_3}(\alpha)$. Darroch and Silvey [43] considered the separability with respect to the likelihood ratio method, which they called "independence". Additionally, Lang and Agresti [45] and Lang [46] discussed the partitioning of goodness of fit statistics. This section describes a property of the statistic for testing the goodness of fit of the S model.

Let n_{ij} denote the observed frequency in cell (i,j) in an $R \times R$ square contingency table with $n = \sum_i \sum_j n_{ij}$. Additionally, let m_{ij} and $\hat{m}_{ij}$ denote the expected frequency in cell (i,j) and the corresponding maximum likelihood estimate (MLE) under a model, respectively. Assume that a multinomial distribution applies to the table. For the S model, MLEs of m_{ij} are given as

$$\hat{m}_{ij} = \begin{cases} \dfrac{n_{ij} + n_{ji}}{2} & (i \neq j), \\ n_{ii} & (i = j). \end{cases}$$

See, for example, Bowker [6] and Bishop et al. [1] (p. 283). Additionally, MLEs of m_{ij} for the CS model are given as

$$\hat{m}_{ij} = \begin{cases} \dfrac{n_U(n_{ij} + n_{ji})}{n_U + n_L} & (i < j), \\ \dfrac{n_L(n_{ij} + n_{ji})}{n_U + n_L} & (i > j), \\ n_{ii} & (i = j), \end{cases}$$

where $n_U = \sum \sum_{i<j} n_{ij}$ and $n_L = \sum \sum_{i>j} n_{ij}$. MLEs of m_{ij} for the GS model are given as

$$\hat{m}_{ij} = \begin{cases} \dfrac{n_{ij}(n_U + n_L)}{2n_U} & (i < j), \\ \dfrac{n_{ij}(n_U + n_L)}{2n_L} & (i > j), \\ n_{ii} & (i = j). \end{cases}$$

It should be noted that the S, CS, and GS models have closed-form estimators. Let $G^2(M)$ denote the likelihood ratio chi-square statistic for testing the goodness of fit of model M. Read [44] noted that

$$G^2(S) = G^2(CS) + G^2(GS). \tag{15}$$

From Equations (9) and (15), $G^2(GS) = G^2(GS|CS)$ (or $G^2(CS) = G^2(CS|GS)$). Therefore, the CS and GS models are separable (i.e., exhibit independence).

Caussinus [8] gave Equation (8). The QS and MH models do not have closed-form estimators. Therefore, it is difficult to obtain a relationship similar to Equation (15). Tomizawa and Tahata [28] proved the relationship between test statistics for the QS and MH models. We review the result briefly.

The QS model is expressed in the log-linear form as

$$\log \pi_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \quad (i = 1, \ldots, R; j = 1, \ldots, R), \tag{16}$$

where $\lambda_{ij}^{XY} = \lambda_{ji}^{XY}$. Without the loss of generality, for example, set $\lambda_R^X = \lambda_R^Y = \lambda_{Rj}^{XY} = \lambda_{iR}^{XY} = 0$. Let $\boldsymbol{\pi} = (\pi_{11}, \ldots, \pi_{1R}, \ldots, \pi_{R1}, \ldots, \pi_{RR})'$. Then the QS model is expressed as

$$\log \boldsymbol{\pi} = \boldsymbol{X}\boldsymbol{\lambda},$$

where $\boldsymbol{\lambda} = (\lambda, \boldsymbol{\lambda}^X, \boldsymbol{\lambda}^Y, \boldsymbol{\lambda}^{XY})'$, $\boldsymbol{\lambda}^X = (\lambda_1^X, \ldots, \lambda_{R-1}^X)$, $\boldsymbol{\lambda}^Y = (\lambda_1^Y, \ldots, \lambda_{R-1}^Y)$, and

$$\boldsymbol{\lambda}^{XY} = (\lambda_{11}^{XY}, \ldots, \lambda_{1,R-1}^{XY}, \lambda_{22}^{XY}, \ldots, \lambda_{2,R-1}^{XY}, \ldots, \lambda_{R-1,R-1}^{XY}).$$

It should be noted that $\boldsymbol{X} = (\boldsymbol{1}_{R^2}, \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_{12})$ is the $R^2 \times K$ matrix with $K = (R^2 + 3R - 2)/2$, where $\boldsymbol{1}_s$ is the $s \times 1$ vector of the 1 element,

$$\boldsymbol{X}_1 = \begin{bmatrix} \boldsymbol{I}_{R-1} \otimes \boldsymbol{1}_R \\ \boldsymbol{O}_{R,R-1} \end{bmatrix}, \quad \boldsymbol{X}_2 = \boldsymbol{1}_R \otimes \begin{bmatrix} \boldsymbol{I}_{R-1} \\ \boldsymbol{0}'_{R-1} \end{bmatrix},$$

and $\boldsymbol{X}_{12}$ is the $R^2 \times R(R-1)/2$ matrix of 1 or 0 elements determined from (16). Additionally, $\boldsymbol{I}_s$ is the $s \times s$ identity matrix, $\boldsymbol{O}_{s,t}$ is the $s \times t$ zero matrix, $\boldsymbol{0}_s$ is the $s \times 1$ zero vector, and $\otimes$ denotes the Kronecker product. The matrix $\boldsymbol{X}$ is a full column rank, which is K. The linear space spanned by the columns of the matrix $\boldsymbol{X}$ is denoted by $S(\boldsymbol{X})$. The dimension of $S(\boldsymbol{X})$ is K and $S(\boldsymbol{X})$ is a subspace of $\mathbb{R}^{R^2}$. Let $S(\boldsymbol{X})^\perp$ denote the orthogonal complement of $S(\boldsymbol{X})$. Let $\boldsymbol{U}$ 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 columns of $\boldsymbol{U}$ span $S(\boldsymbol{X})^\perp$. Since $\boldsymbol{U}'\boldsymbol{X} = \boldsymbol{O}_{d_1,K}$, the QS model can be expressed as $h_1(\boldsymbol{\pi}) = \boldsymbol{0}_{d_1}$, where $h_1(\boldsymbol{\pi}) = \boldsymbol{U}' \log \boldsymbol{\pi}$.

The MH model is defined as

$$\pi_{i\bullet} - \pi_{\bullet i} = 0 \quad (i = 1, \ldots, R-1).$$

Let $h_2(\boldsymbol{\pi}) = \boldsymbol{W}\boldsymbol{\pi}$, where $\boldsymbol{W} = \boldsymbol{X}_1' - \boldsymbol{X}_2'$. The MH model can be expressed as $h_2(\boldsymbol{\pi}) = \boldsymbol{0}_{d_2}$, where $d_2 = R - 1$. The columns of $\boldsymbol{W}'$ belong to subspace $S(\boldsymbol{X})$. It should be noted that $\boldsymbol{W}\boldsymbol{U} = \boldsymbol{O}_{d_2,d_1}$. Since Equation (8) holds, the S model can be expressed as

$$h_3(\boldsymbol{\pi}) = (h_1'(\boldsymbol{\pi}), h_2'(\boldsymbol{\pi}))' = \boldsymbol{0}_{d_3},$$

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

Let $\boldsymbol{p} = (p_{11}, \ldots, p_{1R}, \ldots, p_{R1}, \ldots, p_{RR})'$ denote the sample proportions, where $p_{ij} = n_{ij}/n$. From the multivariate central limit theorem, $\sqrt{n}(\boldsymbol{p} - \boldsymbol{\pi})$ has an asymptotic normal distribution with mean $\boldsymbol{0}_{R^2}$ and covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\pi}) = \boldsymbol{diag}(\boldsymbol{\pi}) - \boldsymbol{\pi}\boldsymbol{\pi}'$, where $\boldsymbol{diag}(\boldsymbol{\pi})$ is the diagonal matrix with the elements of $\boldsymbol{\pi}$ on the main diagonal. In an analogous manner to Bhapkar [63], the Wald statistic can be derived for the S model. Let $\boldsymbol{H}_s(\boldsymbol{\pi}) = \partial h_s(\boldsymbol{\pi})/\partial \boldsymbol{\pi}'$ for $s = 1, 2, 3$. Using the delta method, $\sqrt{n}(h_3(\boldsymbol{p}) - h_3(\boldsymbol{\pi}))$ has an asymptotic normal distribution with mean $\boldsymbol{0}_{d_3}$ and covariance matrix of

$$\begin{bmatrix} \boldsymbol{H}_1(\boldsymbol{\pi})\boldsymbol{\Sigma}(\boldsymbol{\pi})\boldsymbol{H}_1'(\boldsymbol{\pi}) & \boldsymbol{O}_{d_1,d_2} \\ \boldsymbol{O}_{d_2,d_1} & \boldsymbol{H}_2(\boldsymbol{\pi})\boldsymbol{\Sigma}(\boldsymbol{\pi})\boldsymbol{H}_2'(\boldsymbol{\pi}) \end{bmatrix},$$

because $\boldsymbol{H}_1(\boldsymbol{\pi})\boldsymbol{\Sigma}(\boldsymbol{\pi})\boldsymbol{H}_2'(\boldsymbol{\pi}) = \boldsymbol{U}'\boldsymbol{W}' = \boldsymbol{O}_{d_1,d_2}$. Then $T_3 = T_1 + T_2$, where

$$T_s = n h_s'(\boldsymbol{p})[\boldsymbol{H}_s(\boldsymbol{p})\boldsymbol{\Sigma}(\boldsymbol{p})\boldsymbol{H}_s'(\boldsymbol{p})]^{-1} h_s(\boldsymbol{p}).$$

Let $W(M)$ denote the Wald statistic for testing goodness of fit of model M. Since T_1, T_2, and T_3 correspond to $W(QS), W(MH)$, and $W(S)$, respectively, we obtain

$$W(S) = W(QS) + W(MH). \tag{17}$$

Additionally, from the asymptotic equivalence of the Wald statistic and the likelihood ratio statistic (Rao [64], Section 6e. 3), we obtain

$$G^2(S) = G^2(QS) + G^2(MH) + o_p(1). \tag{18}$$

Therefore, the QS and MH models are separable and exhibit asymptotic independence. Caussinus' result (8) shows a good property, such as Equations (17) and (18).

Tahata and Tomizawa [57] proved the asymptotic separability of the LS_k and ME_k models for Equation (12). Tahata [40] extended the result for Equation (13). Additionally, Tahata et al. [61] provided the property of the test statistics related to Equation (14). Separable hypotheses for the S model are discussed elsewhere, for example, by Saigusa et al. [65] and Fujisawa and Tahata [58].

5. Measure of the Departure from Symmetry

As described in Section 1, the analysis of contingency tables is interested in whether row and column variables are independent. If the independence hypothesis is rejected, then the association between the variables is of interest. Various coefficients have been proposed to measure the association such as *gamma*, *Yule's Q*, *Kendall's tau-b*, *Kendall's tau*, and *Somers' d*. See (Bishop et al. [1], Ch. 11) and (Agresti [2], pp. 184–192). On the other hand, Tomizawa [66] proposed two kinds of measures to represent the degree of departure from the S model. Tomizawa et al. [67] gave a generalization of these measures. Their generalization is expressed as the average of the power divergence (Cressie and Read [68]). Additionally, Tomizawa et al. [69], Tahata et al. [70], and Iki et al. [71] have proposed measures to represent the degree of departure from symmetry and the marginal homogeneity. This section reviews a measure to represent the degree of departure from the QS model.

Consider an $R \times R$ square contingency table with *nominal* categories. Caussinus [8] proposed the QS model defined as Equation (1). Using the conditional probability (π_{ij}^c), the QS model can be expressed as

$$Q_{ijk} = Q_{kji} \quad (1 \leq i < j < k \leq R),$$

where $Q_{ijk} = \pi_{ij}^c \pi_{jk}^c \pi_{ki}^c$ and $Q_{kji} = \pi_{kj}^c \pi_{ji}^c \pi_{ik}^c$.

Let

$$\Delta = \sum_{i<j<k} \left(Q_{ijk} + Q_{kji} \right).$$

Additionally, let

$$Q_{ijk}^* = \frac{Q_{ijk}}{\Delta}, \quad Q_{kji}^* = \frac{Q_{kji}}{\Delta}, \quad C_{ijk}^* = C_{kji}^* = \frac{1}{2}(Q_{ijk}^* + Q_{kji}^*),$$

for $1 \leq i < j < k \leq R$. Assuming that $Q_{ijk} + Q_{kji} \neq 0$ for $1 \leq i < j < k \leq R$, Tahata et al. [41] proposed a measure defined as

$$\Phi_{QS}^{(\lambda)} = \frac{1}{2^\lambda - 1} \sum_{i<j<k} \left[Q_{ijk}^* \left\{ \left(\frac{Q_{ijk}^*}{C_{ijk}^*} \right)^\lambda - 1 \right\} + Q_{kji}^* \left\{ \left(\frac{Q_{kji}^*}{C_{kji}^*} \right)^\lambda - 1 \right\} \right] \quad (\lambda > -1).$$

When $\lambda = 0$, this measure is defined by the value taken at the limit as $\lambda \to 0$. Thus

$$\Phi_{QS}^{(0)} = \frac{1}{\log 2} \sum_{i<j<k} \left\{ Q_{ijk}^* \log \left(\frac{Q_{ijk}^*}{C_{ijk}^*} \right) + Q_{kji}^* \log \left(\frac{Q_{kji}^*}{C_{kji}^*} \right) \right\}.$$

$\Phi_{QS}^{(\lambda)}$ is the modification of the power divergence between $\left\{Q_{ijk}^*\right\}$ and $\left\{C_{ijk}^*\right\}$. Especially, $\Phi_{QS}^{(0)}$ is the modification of the Kullback–Leibler divergence between them. It should be noted that the user chooses the real value λ.

In addition, measure $\Phi_{QS}^{(\lambda)}$ can be expressed using the diversity index of degree λ (Patil and Taillie [72]). Let

$$Q_{ijk}^c = \frac{Q_{ijk}}{Q_{ijk} + Q_{kji}}, \quad Q_{kji}^c = \frac{Q_{kji}}{Q_{ijk} + Q_{kji}} \quad (1 \le i < j < k \le R). \tag{19}$$

For paraphrasing Equation (19), please see Tahata and Kozai [73] and Tahata et al. [74]. Then for $\lambda > -1$

$$\Phi_{QS}^{(\lambda)} = \begin{cases} 1 - \dfrac{2^\lambda}{2^\lambda - 1} \displaystyle\sum_{i<j<k} \left(Q_{ijk}^* + Q_{kji}^*\right) \left\{1 - \left(Q_{ijk}^c\right)^{\lambda+1} - \left(Q_{kji}^c\right)^{\lambda+1}\right\} & (\lambda \ne 0), \\[2mm] 1 - \dfrac{1}{\log 2} \displaystyle\sum_{i<j<k} \left(Q_{ijk}^* + Q_{kji}^*\right) \left(-Q_{ijk}^c \log Q_{ijk}^c - Q_{kji}^c \log Q_{kji}^c\right) & (\lambda = 0). \end{cases}$$

Namely, measure $\Phi_{QS}^{(\lambda)}$ is a weighted sum of the diversity index, which includes the Shannon entropy (Shannon [75]) when $\lambda = 0$ and the Gini concentration (Giorgi and Gigliarano [76]) when $\lambda = 1$.

Measure $\Phi_{QS}^{(\lambda)}$ lies between 0 and 1. For any $\lambda > -1$, the value of the measure is 0 if and only if the $R \times R$ contingency table has a QS structure, while its value is 1 if and only if the degree of departure from QS is the largest, in the sense that $Q_{ijk}^c = 0$ (then $Q_{kji}^c = 1$) or $Q_{kji}^c = 0$ (then $Q_{ijk}^c = 1$) for any $1 \le i < j < k \le R$. When the value of the measure is 1, for any $1 \le i < j < k \le R$, then $\pi_{ij}\pi_{jk}\pi_{ki} = 0$ or $\pi_{ji}\pi_{kj}\pi_{ik} = 0$ holds. Namely, for any $1 \le i < j < k \le R$, at least one of π_{ij}^c, π_{jk}^c, and π_{ki}^c is equal to 0, or at least one of π_{ji}^c, π_{kj}^c, and π_{ik}^c is equal to 0. In other words, for any $1 \le i < j < k \le R$, the complete asymmetry arises for at least one pair of symmetric cells. That is, it gives the partial complete asymmetry of cell probabilities. The partial complete asymmetry is a weaker condition than the complete asymmetry of cell probabilities used in Tomizawa [66] and Tomizawa et al. [67]. Additionally, the QS model indicates the symmetry structure of the odds ratio, while $\Phi_{QS}^{(\lambda)} = 1$ indicates the complete asymmetry of the odds ratio because $\theta_{(i<j;j<k)}/\theta_{(j<k;i<j)} = 0$ or ∞ for any $1 \le i < j < k \le R$. It is natural that the degree of departure from the QS model is the largest.

Consider an approximate standard error and large-sample confidence interval for measure $\Phi_{QS}^{(\lambda)}$. Let $\hat{\Phi}_{QS}^{(\lambda)}$ denote the sample version of $\Phi_{QS}^{(\lambda)}$, which is given by $\Phi_{QS}^{(\lambda)}$ when $\boldsymbol{\pi}$ is replaced by $\boldsymbol{p}$. $\hat{\Phi}_{QS}^{(\lambda)}$ has an asymptotic normal distribution with mean $\Phi_{QS}^{(\lambda)}$ and variance $\sigma^2[\hat{\Phi}_{QS}^{(\lambda)}]$ using the delta method (Bishop et al. [1] (Section 14.6)). See Tahata et al. [41] for details of $\sigma^2[\hat{\Phi}_{QS}^{(\lambda)}]$.

Tahata and Kozai [73] proposed a measure to represent the degree of departure from the EQS model proposed by Tomizawa [62]. Additionally, Tahata et al. [74] developed a different measure to represent the degree of departure from the QS model and partitioned their measure into two components. Ando [77] discussed the bivariate index vector to concurrently analyze both the degree and the direction of departure from the QS model.

6. Discussions

When the QS model does not fit a given dataset well, often whether an extended model is suitable is evaluated. Tomizawa [62] proposed the EQS model for a square contingency table with ordinal categories. On the other hand, Tahata [78] proposed the quasi-asymmetry

(QA) model for a square contingency table with nominal categories. The QA model is defined as

$$\log \pi_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \quad (i = 1, \ldots, R; j = 1, \ldots, R),$$

where $\lambda_{ij}^{XY} = \Delta_{ij} + \lambda_{ji}^{XY}$ and $|\Delta_{ij}| = \Delta (\geq 0)$ for $i < j$. Using the odds ratio, the QA model can be expressed as

$$\theta_{(i<j;j<R)} = e^{\Delta_{ij}} \theta_{(j<R;i<j)} \quad (1 \leq i < j \leq R - 1).$$

This states that the odds of the symmetric odds ratio are equal to e^{Δ} for some $i < j$ and $e^{-\Delta}$ for other $i < j$. Using Q_{ijk}^c and Q_{kji}^c in Equation (19), the QA model can be expressed as

$$|Q_{ijR}^c - Q_{Rji}^c| = \Delta^* \quad (1 \leq i < j \leq R - 1),$$

where $\Delta^* = (e^{\Delta} - 1)/(e^{\Delta} + 1)$. When $\Delta = 0$ ($\Delta^* = 0$), the QA model is reduced to the QS model. If the QA model holds, parameter Δ^* is useful to visualize the degree of departure from QS because $0 \leq \Delta^* < 1$. Additionally, the value of Δ^* approaches 1 as the value of Δ increases. The parameter Δ^* may be effective to express the degree of departure from QS.

If a square table has a structure of QA, then measure $\Phi_{QS}^{(\lambda)}$ is simply expressed as

$$\Phi_{QS}^{(\lambda)} = 1 - \frac{\lambda 2^{\lambda}}{(2^{\lambda} - 1) \sum_{s<t<u}(Q_{stu} + Q_{uts})} \sum_{i<j<k} \left(Q_{ijk} + Q_{kji}\right) H_{ijk}^{(\lambda)}(\Delta)$$

where

$$H_{ijk}^{(\lambda)}(\Delta) = \frac{1}{\lambda} \left\{ 1 - \left(\frac{e^{\Delta_{ij}+\Delta_{jk}}}{e^{\Delta_{ij}+\Delta_{jk}} + e^{\Delta_{ik}}}\right)^{\lambda+1} - \left(\frac{e^{\Delta_{ik}}}{e^{\Delta_{ij}+\Delta_{jk}} + e^{\Delta_{ik}}}\right)^{\lambda+1} \right\}.$$

Namely, measure $\Phi_{QS}^{(\lambda)}$ can be expressed as the weighted sum of the function of Δ. Herein, we note that (i) when $\Delta = 0$, $H_{ijk}^{(\lambda)}(0) = (2^{\lambda} - 1)/(\lambda 2^{\lambda})$ for any $i < j < k$. That is, the value of the measure equals 0 (i.e., the QS model holds). (ii) If we take the limit as $\Delta \to \infty$, then $Q_{ijk}^c = 0$ (then $Q_{kji}^c = 1$) or $Q_{kji}^c = 0$ (then $Q_{ijk}^c = 1$) for any $i < j < k$. That is, the value of the measure equals 1. Therefore, the maximum degree of departure from QS can be interpreted as the limitation of QA model with $\Delta \to \infty$.

7. Concluding Remarks

Previously, Tomizawa and Tahata [28] reviewed topics on various symmetry models to analyze square contingency tables, and Tahata and Tomizawa [29] summarized topics related to the symmetry and asymmetry models. This paper mainly reviews four recent developments in the analysis of square contingency tables: (1) modeling based on the *f*-divergence, (2) the necessary and sufficient condition of symmetry, (3) partition of test statistics for symmetry, and (4) the measure of departure from symmetry. These are based on the results in Tahata [40] and Tahata et al. [41].

This paper focuses on the analysis of square contingency tables. Previous papers have examined contingency tables with other structures. For example, Tahata and Tomizawa [79] discussed the double symmetry, which is characterized as both symmetry and point symmetry. In this review, the problem of estimation of cell probabilities and goodness of fit test are omitted. For such issue, Kateri [3], Lawal [80] and Tan [81] would be useful to study computational aspects using R or SAS. Additionally, Lang [82,83] provided the methodology and corresponding R code.

In this paper, the models related to the QS model are described. Especially, the characterization of models and its properties are discussed. The models described this paper is useful for analyzing real dataset. For example, unaided distance vision data, social

mobility data, father–son matched educational level data, longitudinal data in biomedical research, and so on. Such applications can be seen in each of references. For details of the estimation of cell probabilities and goodness of fit test, please see corresponding papers.

Various topics are generalized for multi-way contingency tables with the same classifications. Bhapkar and Darroch [84] proposed the QS and MH models for a general order and gave the generalization of Equation (8). Tomizawa and Tahata [28] proved the generalization of Equation (18). For multi-way contingency tables, the structures of asymmetries are considered in Tahata and Tomizawa [57], Tahata et al. [85], Tahata et al. [86] and Shinoda et al. [87]. Moreover, Tahata and Tomizawa [88] and Yoshimoto et al. [59] treated the problem of point symmetry for multi-way contingency tables.

In future work, the generalization of the results in described in this review will be considered for the multi-way contingency tables. Additionally, the goodness of fit test for symmetry and asymmetry models should be considered for sparse contingency tables.

References

1. Bishop, Y.M.; Fienberg, S.E.; Holland, P.W. *Discrete Multivariate Analysis: Theory and Practice*; The MIT Press: Cambridge, MA, USA, 1975.
2. Agresti, A. *Analysis of Ordinal Categorical Data*; John Wiley: Hoboken, NJ, USA, 2010.
3. Kateri, M. *Contingency Table Analysis. Methods and Implementation Using R*; Birkhäuser: Basel, Switzerland, 2014.
4. Kateri, M. ϕ-divergence in contingency table analysis. *Entropy* **2018**, *20*, 324. [CrossRef] [PubMed]
5. Fujisawa, K.; Tahata, K. Quasi Association Models for Square Contingency Tables with Ordinal Categories. *Symmetry* **2022**, *14*, 805. [CrossRef]
6. Bowker, A.H. A test for symmetry in contingency tables. *J. Am. Stat. Assoc.* **1948**, *43*, 572–574. [CrossRef] [PubMed]
7. Stuart, A. A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika* **1955**, *42*, 412–416. [CrossRef]
8. Caussinus, H. Contribution à l'analyse statistique des tableaux de corrélation. *Ann. Fac. Sci. Toulouse Math.* **1965**, *29*, 77–183. [CrossRef]
9. Agresti, A. Links between binary and multi-category logit item response models and quasi-symmetric loglinear models. *Ann. Fac. Sci. Toulouse Math.* **2002**, *11*, 443–454. [CrossRef]
10. Goodman, L.A. Contributions to the statistical analysis of contingency tables: Notes on quasi-symmetry, quasi-independence, log-linear models, log-bilinear models, and correspondence analysis models. *Ann. Fac. Sci. Toulouse Math.* **2002**, *11*, 525–540. [CrossRef]
11. McCullagh, P. Quasi-symmetry and representation theory. *Ann. Fac. Sci. Toulouse Math.* **2002**, *11*, 541–561. [CrossRef]
12. Fienberg, S.E.; van der Heijden, P.G.M. Introduction to special issue on quasi-symmetry and categorical data analysis. *Ann. Fac. Sci. Toulouse Math.* **2002**, *11*, 439–441. [CrossRef]
13. Bergsma, W.P.; Rudas, T. Modeling conditional and marginal association in contingency tables. *Ann. Fac. Sci. Toulouse Math.* **2002**, *11*, 455–468. [CrossRef]
14. Dossou-Gbété, S.; Grorud, A. Biplots for matched two-way tables. *Ann. Fac. Sci. Toulouse Math.* **2002**, *11*, 469–483. [CrossRef]
15. Erosheva, E.A.; Fienberg, S.E.; Junker, B.W. Alternative statistical models and representations for large sparse multi-dimensional contingency tables. *Ann. Fac. Sci. Toulouse Math.* **2002**, *11*, 485–505. [CrossRef]
16. De Falguerolles, A.; van der Heijden, P.G.M. Reduced rank quasi-symmetry and quasi-skew symmetry: A generalized bi-linear model approach. *Ann. Fac. Sci. Toulouse Math.* **2002**, *11*, 507–524. [CrossRef]
17. Stigler, S. The missing early history of contingency tables. *Ann. Fac. Sci. Toulouse Math.* **2002**, *11*, 563–573. [CrossRef]
18. Thélot, C. L'analyse statistique des tables de mobilité à l'aide du modèle quasisymétrique et de ses dérivés. *Ann. Fac. Sci. Toulouse Math.* **2002**, *11*, 575–585. [CrossRef]
19. Caussinus, H. Some concluding observations. *Ann. Fac. Sci. Toulouse Math.* **2002**, *11*, 587–591. [CrossRef]
20. Tomizawa, S. Diagonals-parameter symmetry model for cumulative probabilities in square contingency tables with ordered categories. *Biometrics* **1993**, *49*, 883–887. [CrossRef]

21. Miyamoto, N.; Ohtsuka, W.; Tomizawa, S. Linear diagonals-parameter symmetry and quasi-symmetry models for cumulative probabilities in square contingency tables with ordered categories. *Biom. J.* **2004**, *46*, 664–674. [CrossRef]
22. Kateri, M.; Gottard, A.; Tarantola, C. Generalised quasi-symmetry models for ordinal contingency tables. *Aust. N. Z. J. Stat.* **2017**, *59*, 239–253. [CrossRef]
23. Booth, J.G.; Capanu, M.; Heigenhauser, L. Exact conditional P value calculation for the quasi-symmetry model. *J. Comput. Graph. Stat.* **2005**, *14*, 716–725. [CrossRef]
24. Krampe, A.; Kateri, M.; Kuhnt, S. Asymmetry models for square contingency tables: Exact tests via algebraic statistics. *Stat. Comput.* **2011**, *21*, 55–67. [CrossRef]
25. Rapallo, F. Algebraic markov bases and MCMC for two-way contingency tables. *Scand. J. Stat.* **2003**, *30*, 385–397. [CrossRef]
26. Pardo, L.; Martin, N. Minimum phi-divergence estimators and phi-divergence test statistics in contingency tables with symmetry structure: An overview. *Symmetry* **2010**, *2*, 1108–1120. [CrossRef]
27. Gottard, A.; Marchetti, G.M.; Agresti, A. Quasi-symmetric graphical log-linear models. *Scand. J. Stat.* **2011**, *38*, 447–465. [CrossRef]
28. Tomizawa, S.; Tahata, K. The analysis of symmetry and asymmetry: Orthogonality of decomposition of symmetry into quasi-symmetry and marginal symmetry for multi-way tables. *J. Soc. Fr. Stat.* **2007**, *148*, 3–36.
29. Tahata, K.; Tomizawa, S. Symmetry and asymmetry models and decompositions of models for contingency tables. *SUT J. Math.* **2014**, *50*, 131–165.
30. Bocci, C.; Rapallo, F. Exact tests to compare contingency tables under quasi independence and quasi-symmetry. *J. Algebr. Stat.* **2019**, *10*, 13–29. [CrossRef]
31. Khan, Z.A.; Tewari, R.C. Markov reversibility, quasi-symmetry, and marginal homogeneity in cyclothymiacs geological successions. *Int. J. Geoinform. Geol. Sci.* **2021**, *8*, 9–25.
32. Altun, G. Quasi local odds symmetry model for square contingency table with ordinal categories. *J. Stat. Comput. Simul.* **2019**, *89*, 2899–2913. [CrossRef]
33. Tahata, K.; Ochiai, T.; Matsushima, U. Asymmetry models and model selection in square contingency tables with ordinal categories. In Proceedings of the 2020 International Symposium on Information Theory and Its Applications (ISITA), Kapolei, HI, USA, 24–27 October 2020; pp. 573–577.
34. Karadağ, O.; Altun, G.; Aktaş, S. Assessment of SNP-SNP interactions by using square contingency table analysis. *Ann. Braz. Acad. Sci.* **2020**, *92*, e20190465. [CrossRef]
35. Altunay, S.A.; Yilmaz, A.E. Median Distance Model for Likert-Type Items in Contingency Table Analysis. Accepted-November 2021. Available online: https://revstat.ine.pt/index.php/REVSTAT/article/view/401 (accessed on 1 February 2022).
36. Ando, S. Asymmetry models based on ordered score and separations of symmetry model for square contingency tables. *Biom. Lett.* **2021**, *58*, 27–39. [CrossRef]
37. Ando, S. Orthogonal decomposition of the sum-symmetry model for square contingency tables with ordinal categories: Use of the exponential sum-symmetry model. *Biom. Lett.* **2021**, *58*, 95–104. [CrossRef]
38. Ando, S. Odds-symmetry model for cumulative probabilities and decomposition of a conditional symmetry model in square contingency tables. *Aust. N. Z. J. Stat.* **2021**, *63*, 674–684. [CrossRef]
39. Ando, S. Asymmetry models based on non-integer scores for square contingency tables. *J. Stat. Theory Appl.* **2022**, *21*, 21–30. [CrossRef]
40. Tahata, K. Separation of symmetry for square tables with ordinal categorical data. *Jpn. J. Stat. Data Sci.* **2020**, *3*, 469–484. [CrossRef]
41. Tahata, K.; Miyamoto, N.; Tomizawa, S. Measure of departure from quasi-symmetry and bradley-terry models for square contingency tables with nominal categories. *J. Korean Stat. Soc.* **2004**, *33*, 129–147.
42. Aitchison, J. Large-sample restricted parametric tests. *J. R. Stat. Soc. Ser.-Stat. Methodol.* **1962**, *24*, 234–250. [CrossRef]
43. Darroch, J.N.; Silvey, S.D. On testing more than one hypothesis. *Ann. Math. Stat.* **1963**, *34*, 555–567. [CrossRef]
44. Read, C.B. Partitioning chi-squape in contingency tables: A teaching approach. *Commun. Stat.-Theory Methods* **1977**, *6*, 553–562. [CrossRef]
45. Lang, J.B.; Agresti, A. Simultaneously modeling joint and marginal distributions of multivariate categorical responses. *J. Am. Stat. Assoc.* **1994**, *89*, 625–632. [CrossRef]
46. Lang, J.B. On the partitioning of goodness-of-fit statistics for multivariate categorical response models. *J. Am. Stat. Assoc.* **1996**, *91*, 1017–1023. [CrossRef]
47. McCullagh, P. A class of parametric models for the analysis of square contingency tables with ordered categories. *Biometrika* **1978**, *65*, 413–418. [CrossRef]
48. Agresti, A. A simple diagonals-parameter symmetry and quasi-symmetry model. *Stat. Probab. Lett.* **1983**, *1*, 313–316. [CrossRef]
49. Tomizawa, S. An extended linear diagonals-parameter symmetry model for square contingency tables with ordered categories. *Metron* **1991**, *49*, 401–409.
50. Ireland, C.T.; Ku, H.H.; Kullback, S. Symmetry and Marginal Homogeneity of an $r \times r$ Contingency Table. *J. Am. Stat. Assoc.* **1969**, *64*, 1323–1341. [CrossRef]
51. Gilula, Z.; Krieger, A.M.; Ritov, Y. Ordinal Association in Contingency Tables: Some Interpretive Aspects. *J. Am. Stat. Assoc.* **1988**, *83*, 540–545. [CrossRef]
52. Kateri, M.; Papaioannou, T. Asymmetry models for contingency tables. *J. Am. Stat. Assoc.* **1997**, *92*, 1124–1131. [CrossRef]

53. Csiszár, I.; Shields, P.C. *Information Theory and Statistics: A Tutorial*; Now Publishers Inc.: Hanover, Germany, 2004.

54. Kullback, S.; Leibler, R.A. On Information and Sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [CrossRef]

55. Kateri, M.; Agresti, A. A class of ordinal quasi-symmetry models for square contingency tables. *Stat. Probab. Lett.* **2007**, *77*, 598–603. [CrossRef]

56. Kateri, M. Families of generalized quasisymmetry models: A ϕ-divergence approach. *Symmetry* **2021**, *13*, 2297. [CrossRef]

57. Tahata, K.; Tomizawa, S. Generalized linear asymmetry model and decomposition of symmetry for multiway contingency tables. *J. Biom. Biostat.* **2011**, *2*, 1–6. [CrossRef]

58. Fujisawa, K.; Tahata, K. Asymmetry model based on f-divergence and orthogonal decomposition of symmetry for square contingency tables with ordinal categories. *SUT J. Math.* **2020**, *56*, 39–53.

59. Yoshimoto, T.; Tahata, K.; Saigusa, Y.; Tomizawa, S. Quasi point-symmetry models based on f-divergence and decomposition of point-symmetry for multi-way contingency tables. *SUT J. Math.* **2019**, *55*, 103–131.

60. Tahata, K.; Tomizawa, S. Generalized marginal homogeneity model and its relation to marginal equimoments for square contingency tables with ordered categories. *Adv. Data Anal. Classif.* **2008**, *2*, 295–311. [CrossRef]

61. Tahata, K.; Naganawa, M.; Tomizawa, S. Extended linear asymmetry model and separation of symmetry for square contingency tables. *J. Jpn. Stat. Soc.* **2016**, *46*, 189–202. [CrossRef]

62. Tomizawa, S. Three kinds of decompositions for the conditional symmetry model in a square contingency table. *J. Jpn. Stat. Soc.* **1984**, *14*, 35–42.

63. Bhapkar, V.P. A note on the equivalence of two test criteria for hypotheses in categorical data. *J. Am. Stat. Assoc.* **1966**, *61*, 228–235. [CrossRef]

64. Rao, C.R. *Linear Statistical Inference and Its Applications*, 2nd ed.; Wiley: New York, NY, USA, 1973.

65. Saigusa, Y.; Tahata, K.; Tomizawa, S. Orthogonal decomposition of symmetry model using the ordinal quasi-symmetry model based on f-divergence for square contingency tables. *Stat. Probab. Lett.* **2015**, *101*, 33–37. [CrossRef]

66. Tomizawa, S. Two kinds of measures of departure from symmetry in square contingency tables having nominal categories. *Stat. Sin.* **1994**, *4*, 325–334.

67. Tomizawa, S.; Seo, T.; Yamamoto, H. Power-divergence-type measure of departure from symmetry for square contingency tables that have nominal categories. *J. Appl. Stat.* **1998**, *25*, 387–398. [CrossRef]

68. Cressie, N.; Read, T.R.C. Multinomial goodness-of-fit tests. *J. R. Stat. Soc. Ser. B-Methodol.* **1984**, *46*, 440–464. [CrossRef]

69. Tomizawa, S.; Miyamoto, N.; Hatanaka, Y. Measure of asymmetry for square contingency tables having ordered categories. *Aust. N. Z. J. Stat.* **2001**, *43*, 335–349. [CrossRef]

70. Tahata, K.; Iwashita, T.; Tomizawa, S. Measure of departure from conditional marginal homogeneity for square contingency tables with ordered categories. *Statistics* **2008**, *42*, 453–466. [CrossRef]

71. Iki, K.; Tahata, K.; Tomizawa, S. Measure of departure from marginal homogeneity using marginal odds for multi-way tables with ordered categories. *J. Appl. Stat.* **2012**, *39*, 279–295. [CrossRef]

72. Patil, G.P.; Taillie, C. Diversity as a concept and its measurement. *J. Am. Stat. Assoc.* **1982**, *77*, 548–561. [CrossRef]

73. Tahata, K.; Kozai, K. Measuring degree of departure from extended quasi-symmetry for square contingency tables. *Rev. Colomb. Estad.* **2012**, *35*, 55–65.

74. Tahata, K.; Kozai, K.; Tomizawa, S. Partitioning measure of quasi-symmetry for square contingency tables. *Braz. J. Probab. Stat.* **2014**, *28*, 353–366. [CrossRef]

75. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]

76. Giorgi, G.M.; Gigliarano, C. The Gini concentration index: A review of the inference literature. *J. Econ. Surv.* **2017**, *31*, 1130–1148. [CrossRef]

77. Ando, S. A bivariate index vector to measure departure from quasi-symmetry for ordinal square contingency tables. *Austrian J. Stat.* **2021**, *50*, 115–126. [CrossRef]

78. Tahata, K. Quasi-asymmetry model for square tables with nominal categories. *J. Appl. Stat.* **2012**, *39*, 723–729. [CrossRef]

79. Tahata, K.; Tomizawa, S. Double linear diagonals-parameter symmetry and decomposition of double symmetry for square tables. *Stat. Methods Appl.* **2010**, *19*, 307–318. [CrossRef]

80. Lawal, H.B. Using a GLM to decompose the symmetry model in square contingency tables with ordered categories. *J. Appl. Stat.* **2004**, *31*, 279–303. [CrossRef]

81. Tan, T. *Doubly Classified Model with R*; Springer: Singapore, 2017.

82. Lang, J.B. Multinomial-Poisson homogeneous models for contingency tables. *Ann. Stat.* **2004**, *32*, 340–383. [CrossRef]

83. Lang, J.B. Homogeneous Linear Predictor Models for Contingency Tables. *J. Am. Stat. Assoc.* **2005**, *100*, 121–134. [CrossRef]

84. Bhapkar, V.P.; Darroch, J.N. Marginal symmetry and quasi symmetry of general order. *J. Multivar. Anal.* **1990**, *34*, 173–184. [CrossRef]

85. Tahata, K.; Yamamoto, H.; Tomizawa, S. Orthogonality of decompositions of symmetry into extended symmetry and marginal equimoment for multi-way tables with ordered categories. *Austrian J. Stat.* **2008**, *37*, 185–194. [CrossRef]

86. Tahata, K.; Yamamoto, H.; Tomizawa, S. Linear ordinal quasi-symmetry model and decomposition of symmetry for multi-way tables. *Math. Methods Stat.* **2011**, *20*, 158–164. [CrossRef]

87. Shinoda, S.; Tahata, K.; Yamamoto, K.; Tomizawa, S. Marginal continuation odds ratio model and decomposition of marginal homogeneity model for multi-way contingency tables. *Sankhya B* **2021**, *83*, 304–324. [CrossRef]
88. Tahata, K.; Tomizawa, S. Orthogonal decomposition of point-symmetry for multiway tables. *AStA Adv. Stat. Anal.* **2008**, *92*, 255–269. [CrossRef]