SUT Journal of Mathematics
Vol. 55, No. 2 (2019), 109–137

Quasi point-symmetry models based on f-divergence and decomposition of point-symmetry for multi-way contingency tables

Takuya Yoshimoto, Kouji Tahata, Yusuke Saigusa and Sadao Tomizawa

(Received April 1, 2019; Revised July 7, 2019)

Abstract. For two-way contingency tables, Tomizawa (1985) considered the quasi point-symmetry (QP) model and the marginal point-symmetry (MP) model, and gave the theorem that the point-symmetry (PS) model holds if and only if both the QP and MP models hold. Tahata and Tomizawa (2008) provided similar theorems for multi-way tables. For multi-way tables, the present paper proposes the quasi point-symmetry (QP[f]) model based on f-divergence. The QP[f] model includes the QP model in a special case. It also gives the theorem that the PS model holds if and only if both the QP[f] and MP models hold, and the theorem that the test statistic for goodness-of-fit of the PS model is asymptotically equivalent to the sum of those for the decomposed models under the PS model.

AMS 2010 *Mathematics Subject Classification.* 62H17.

Key words and phrases. Marginal point-symmetry, orthogonality, square contingency table.

§1. Introduction

For an $r \times r$ contingency table, let p_{ij} denote the probability that an observation will fall in the ith row and jth column of the table ($i = 1, \ldots, r;\ j = 1, \ldots, r$). The point-symmetry (PS) model is defined by

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

where $i^* = r + 1 - i$ (Wall and Lienert, 1976; Tomizawa, 1985). This model states that the probability that an observation will fall in cell (i, j) is equal to the probability that it falls in point symmetric cell (i^*, j^*) with respect to the center point (or cell). The PS model may be expressed in a log-linear form

$$\log p_{ij} = u_{1(i)} + u_{2(j)} + u_{12(ij)} \quad (i = 1, \ldots, r;\ j = 1, \ldots, r),$$

where $u_{1(i)} = u_{1(i^*)}$, $u_{2(j)} = u_{2(j^*)}$, and $u_{12(ij)} = u_{12(i^*j^*)}$. Tomizawa (1985) considered the quasi point-symmetry (QP) model and the marginal point-symmetry (MP) model. The QP model is defined by

$$\log p_{ij} = u_{1(i)} + u_{2(j)} + u_{12(ij)} \quad (i = 1, \ldots, r;\ j = 1, \ldots, r),$$

where $u_{12(ij)} = u_{12(i^*j^*)}$. A special case of the QP model obtained by putting $\{u_{1(i)} = u_{1(i^*)}\}$ and $\{u_{2(j)} = u_{2(j^*)}\}$ is the PS model. Denote the odds ratio for rows i and j $(> i)$, and columns k and l $(> k)$ by $\theta_{(i<j;k<l)}$. Namely $\theta_{(i<j;k<l)} = (p_{ik}p_{jl})/(p_{jk}p_{il})$. Using odds ratios, the QP model is also expressed as

$$\theta_{(i<j;k<l)} = \theta_{(j^*<i^*;l^*<k^*)} \quad (i < j;\ k < l).$$

Therefore the QP model has its characterization in terms of point-symmetry of odds ratios. The MP model is defined by

$$p_{i\cdot} = p_{i^*\cdot} \quad (i = 1, \ldots, r) \quad \text{and} \quad p_{\cdot j} = p_{\cdot j^*} \quad (j = 1, \ldots, r),$$

where $p_{i\cdot} = \sum_{t=1}^{r} p_{it}$ and $p_{\cdot j} = \sum_{s=1}^{r} p_{sj}$. This indicates that the row (column) marginal distributions are point symmetric with respect to the midpoint of the row (column) categories.

Kateri and Papaioannou (1997), Kateri and Agresti (2007), and Saigusa, Tahata and Tomizawa (2015) have described some models of symmetry based on f-divergence, and considered the property of the model in the information theoretic sense. Also see, for example, Gilula, Krieger and Ritov (1988) and Kateri (2018). Tomizawa (1985) gave the theorem that the PS model holds if and only if both the QP and MP models hold, and Tahata and Tomizawa (2008) showed the decomposition of the PS model for multi-way tables.

In the present paper, we propose the QP model based on f-divergence, and give a theorem that the PS model holds if and only if both the proposed model and the MP model hold for multi-way tables. The proposed model and decomposition theorem include the results of Tomizawa (1985) and Tahata and Tomizawa (2008) as a special case.

Section 2 proposes the new model and the decomposition of the PS model for two-way tables. Section 3 proposes the extended model and decomposition of the PS model for three-way tables. Section 4 proposes the generalized model for multi-way tables and provides a decomposition. Section 5 presents properties of test statistics for goodness-of-fit of models, and Section 6 shows examples. Finally, Section 7 concludes this paper.

§2. Case of two-way tables

2.1. Model

Let $p = (p_{ij})$ and $q = (q_{ij})$ be two discrete finite bivariate probability distributions. The f-divergence between p and q is defined as

$$I^C(p : q) = \sum_i \sum_j q_{ij} f\left(\frac{p_{ij}}{q_{ij}}\right),$$

where f is a convex function on $(0, +\infty)$ with $f(1) = 0$. Also, we take $f(0) = \lim_{u\to 0} f(u)$, $0f(0/0) = 0$, and $0f(a/0) = a \lim_{u\to\infty} [f(u)/u]$ (Csiszár and Shields, 2004). Let f be a twice-differentiable and strictly convex function, and let $F(u) = df(u)/du$. Let $p^{PS}_{ij} = (p_{ij} + p_{i^*j^*})/2$ for $i = 1, \ldots, r$ and $j = 1, \ldots, r$.

We propose the quasi point-symmetry (QP[f]) model based on f-divergence as follows:

$$p_{ij} = p^{PS}_{ij} F^{-1}(u_{1(i)} + u_{2(j)} + u_{12(ij)}) \quad (i = 1, \ldots, r;\ j = 1, \ldots, r),$$

where $u_{12(ij)} = u_{12(i^*j^*)}$. Note that the parameters of the QP[f] model must satisfy the following equation from the relation $p^{PS}_{ij} = (p_{ij} + p_{i^*j^*})/2$,

$$F^{-1}\left(u_{1(i)} + u_{2(j)} + u_{12(ij)}\right) + F^{-1}\left(u_{1(i^*)} + u_{2(j^*)} + u_{12(i^*j^*)}\right) = 2.$$

The QP[f] model can also be expressed as

$$\theta^{[f]}_{(i<j;k<l)} = \theta^{[f]}_{(j^*<i^*;l^*<k^*)} \quad (i < j;\ k < l), \tag{2.1}$$

where

$$\theta^{[f]}_{(i<j;k<l)} = F\left(\frac{p_{ik}}{p^{PS}_{ik}}\right) + F\left(\frac{p_{jl}}{p^{PS}_{jl}}\right) - F\left(\frac{p_{jk}}{p^{PS}_{jk}}\right) - F\left(\frac{p_{il}}{p^{PS}_{il}}\right).$$

Therefore the QP[f] model has its characterization in terms of point-symmetry of $\theta^{[f]}_{(i<j;k<l)}$. We note that $\theta^{[f]}_{(i<j;k<l)}$ with $f(u) = u \log u$, $u > 0$ equals $\log(p_{ik}p_{jl}/p_{jk}p_{il}) - \log\left(p^{PS}_{ik}p^{PS}_{jl}/p^{PS}_{jk}p^{PS}_{il}\right)$. Hence, equation (2.1) reduces to the point-symmetry of odds ratios when we set $f(u) = u \log u$. This enables us to see that the QP[f] model is a generalization of the structure of quasi point-symmetry including the QP model.

The QP[f] model indicates that the monotonic function of conditional point symmetric probability, i.e., $F\left(p_{ij}/p^{PS}_{ij}\right)$, is expressed as the linear function of the row effect term $u_{1(i)}$, the column effect term $u_{2(j)}$, and the association term $u_{12(ij)}$, where the association term has the structure of point-symmetry, i.e., $u_{12(ij)} = u_{12(i^*j^*)}$. Especially, when the function F is a log-function, the QP[f] model is the QP model, which is also expressed as the structure of point-symmetry of odds ratios by eliminating the row effect and the column effect. So, equation (2.1) may be interpreted as the structure of point-symmetry of, so to speak, generalized odds ratio based on function F (including a log-function in a special case) by eliminating the row effect and the column effect. The PS model indicates that each of the row effect term, the column effect term, and the association term has the structure of point-symmetry (see Section 1). When the PS model does not hold, we are interested in how the association term has the structure of point-symmetry, i.e., in seeing the function F satisfying equation (2.1) because the function F satisfying equation (2.1) may not be a log-function (i.e., the QP model). We obtain the following theorem.

Theorem 2.1. *The QP[f] model for $\{p_{ij}\}$ minimizes the f-divergence between $\{p_{ij}\}$ and $\left\{p_{ij}^{PS}\right\}$ with the structure of the PS model under the condition that row marginals $\{p_{i\cdot}\}$, column marginals $\{p_{\cdot j}\}$ and sums $\{p_{ij} + p_{i^*j^*}\}$ are given.*

Proof. Let $p^{PS} = \left(p_{ij}^{PS}\right)$. Consider a probability distribution p minimizes the f-divergence $I^C(p : p^{PS})$ under the restrictions of

$$p_{k\cdot} = p_{k0} \quad (k = 1, \dots, r), \quad p_{\cdot l} = p_{0l} \quad (l = 1, \dots, r),$$

and

$$p_{kl} + p_{k^*l^*} = 2p_{kl}^{(00)} \quad (k = 1, \dots, r;\ l = 1, \dots, r),$$

where $p_{kl}^{(00)} = p_{k^*l^*}^{(00)}$, and the values p_{k0}, p_{0l} and $p_{kl}^{(00)}$ are constants. Consider the Lagrange function as

$$L(p_{ij}) = I^C(p : p^{PS}) + \sum_{k=1}^{r} \lambda_{1(k)}(p_{k\cdot} - p_{k0}) + \sum_{l=1}^{r} \lambda_{2(l)}(p_{\cdot l} - p_{0l}) + \frac{1}{2}\sum_{k=1}^{r}\sum_{l=1}^{r} \lambda_{12(kl)}(p_{kl} + p_{k^*l^*} - 2p_{kl}^{(00)}),$$

where $\lambda_{12(kl)} = \lambda_{12(k^*l^*)}$. Equating the derivation of $L(p_{ij})$ to 0 with respect to p_{ij}, $\lambda_{1(i)}$, $\lambda_{2(j)}$ and $\lambda_{12(ij)}$ gives

$$F\left(\frac{p_{ij}}{p_{ij}^{PS}}\right) + \lambda_{1(i)} + \lambda_{2(j)} + \lambda_{12(ij)} = 0 \quad (i = 1, \dots, r;\ j = 1, \dots, r),$$

$$p_{i\cdot} - p_{i0} = 0 \quad (i = 1, \dots, r),$$

$$p_{\cdot j} - p_{0j} = 0 \quad (j = 1, \dots, r),$$

$$p_{ij} + p_{i^*j^*} - 2p_{ij}^{(00)} = 0 \quad (i = 1, \ldots, r;\ j = 1, \ldots, r).$$

Hence, we obtain

$$\frac{p_{ij}}{p_{ij}^{PS}} = F^{-1}\left(-\lambda_{1(i)} - \lambda_{2(j)} - \lambda_{12(ij)}\right).$$

Denoting $-\lambda_{1(i)}$, $-\lambda_{2(j)}$ and $-\lambda_{12(ij)}$ by $u_{1(i)}$, $u_{2(j)}$ and $u_{12(ij)}$, respectively, we have

$$p_{ij} = p_{ij}^{PS} F^{-1}(u_{1(i)} + u_{2(j)} + u_{12(ij)}), \tag{2.2}$$

where $u_{12(ij)} = u_{12(i^*j^*)}$. Also $\{u_{1(i)}\}$, $\{u_{2(j)}\}$ and $\{u_{12(ij)}\}$ are expressed as functions of $\{p_{i0}\}$, $\{p_{0j}\}$ and $\left\{p_{ij}^{(00)}\right\}$. Since $\{p_{i0}\}$, $\{p_{0j}\}$ and $\left\{p_{ij}^{(00)}\right\}$ are arbitrary values, equation (2.2) is the QP$[f]$ model. The proof is complete. □

When we set $f(u) = u \log u$, $u > 0$, the QP$[f]$ model is expressed as

$$p_{ij} = p_{ij}^{PS} \exp(u_{1(i)} + u_{2(j)} + u_{12(ij)} - 1) \quad (i = 1, \ldots, r;\ j = 1, \ldots, r),$$

where $u_{12(ij)} = u_{12(i^*j^*)}$. This may be expressed as

$$p_{ij} = p_{ij}^{PS} \frac{2a_i b_j}{1 + a_i b_j} \quad (i = 1, \ldots, r;\ j = 1, \ldots, r),$$

with $a_i = \exp(u_{1(i)} - u_{1(i^*)})$ and $b_j = \exp(u_{2(j)} - u_{2(j^*)})$. This equation is equivalent to

$$\frac{p_{ij}}{p_{i^*j^*}} = a_i b_j \quad (i = 1, \ldots, r;\ j = 1, \ldots, r).$$

From (2.1), the QP$[f]$ model with $f(u) = u \log u$ can be expressed as

$$\theta_{(i<j;k<l)} = \theta_{(j^*<i^*;l^*<k^*)} \quad (i < j;\ k < l).$$

Therefore the QP$[f]$ model with $f(u) = u \log u$ is equivalent to the QP model. In addition, we can see from Theorem 2.1 that the QP model is the closest model to the PS model in terms of the f-divergence with $f(u) = u \log u$, i.e., Kullback-Leibler distance, under the conditions that row marginals $\{p_{i\cdot}\}$, column marginals $\{p_{\cdot j}\}$ and sums $\{p_{ij} + p_{i^*j^*}\}$ are given.

Next, when $f(u) = (1 - u)^2$, $u > 0$, the QP$[f]$ model reduces to

$$p_{ij} = p_{ij}^{PS} \left(\frac{u_{1(i)} + u_{2(j)} + u_{12(ij)}}{2} + 1\right) \quad (i = 1, \ldots, r;\ j = 1, \ldots, r),$$

where $u_{12(ij)} = u_{12(i^*j^*)}$. We shall refer to this model as the Pearsonian-QP model, which is the closest model to the PS model when divergence is measured by the Pearsonian distance. It can easily be verified that $u_{12(ij)} = -(u_{1(i)} + u_{1(i^*)} + u_{2(j)} + u_{2(j^*)})/2$ for $i = 1, \ldots, r$ and $j = 1, \ldots, r$. Thus the Pearsonian-QP model becomes

$$p_{ij} = p_{ij}^{PS}(a_i + b_j + 1) \quad (i = 1, \ldots, r;\ j = 1, \ldots, r),$$

with $a_i = (u_{1(i)} - u_{1(i^*)})/4$ and $b_j = (u_{2(j)} - u_{2(j^*)})/4$. Hence, this model may be expressed as

$$\frac{p_{ij}}{p_{i^*j^*}} = \frac{1 + a_i + b_j}{1 - (a_i + b_j)} \quad (i = 1, \ldots, r;\ j = 1, \ldots, r).$$

From (2.1), the Pearsonian-QP model can be expressed as

$$\theta^P_{(i<j;k<l)} = \theta^P_{(j^*<i^*;l^*<k^*)} \quad (i < j;\ k < l),$$

where

$$\theta^P_{(i<j;k<l)} = 2\left(\frac{p_{ik}}{p_{ik}^{PS}} + \frac{p_{jl}}{p_{jl}^{PS}} - \frac{p_{jk}}{p_{jk}^{PS}} - \frac{p_{il}}{p_{il}^{PS}}\right).$$

Therefore the Pearsonian-QP model has its characterization in terms of point-symmetry of $\theta^P_{(i<j;k<l)}$.

Read and Cressie (1988) considered the power divergence for p and q, which is defined by,

$$I^\lambda(p:q) = \frac{1}{\lambda(\lambda+1)} \sum_i \sum_j p_{ij}\left[\left(\frac{p_{ij}}{q_{ij}}\right)^\lambda - 1\right] \quad (-\infty < \lambda < \infty),$$

where the values at $\lambda = 0$ and $\lambda = -1$ are the continuous limits as $\lambda \to 0$ and $\lambda \to -1$, respectively. The power divergence is a special case of the f-divergence because $I^C(p : q)$ with $f^\lambda(u) = (\lambda(\lambda+1))^{-1}(u^{\lambda+1} - u)$, $u > 0$, reduces to the power divergence. The QP[f] model with $f^\lambda(u)$ becomes

$$p_{ij} = p_{ij}^{PS}\left[\lambda(u_{1(i)} + u_{2(j)} + u_{12(ij)}) + \frac{1}{\lambda+1}\right]^{\frac{1}{\lambda}} \quad (i = 1, \ldots, r;\ j = 1, \ldots, r),$$

where $u_{12(ij)} = u_{12(i^*j^*)}$. Note that $I^0(p : q)$ and $I^1(p : q)$ are the Kullback-Leibler distance and the Pearsonian distance multiplied by $1/2$, respectively. For the power divergence, we can see that the QP[f] model has its characterization in terms of point-symmetry of $\theta^{(\lambda)}_{(i<j;k<l)}$, where

$$\theta^{(\lambda)}_{(i<j;k<l)} = \frac{1}{\lambda}\left[\left(\frac{p_{ik}}{p_{ik}^{PS}}\right)^\lambda + \left(\frac{p_{jl}}{p_{jl}^{PS}}\right)^\lambda - \left(\frac{p_{jk}}{p_{jk}^{PS}}\right)^\lambda - \left(\frac{p_{il}}{p_{il}^{PS}}\right)^\lambda\right].$$

2.2. Decomposition of point-symmetry

Tomizawa (1985) showed the decomposition of the PS model into the QP and MP models. We obtain the following theorem, which includes the result of Tomizawa (1985) in a special case.

Theorem 2.2. *The PS model holds if and only if both the QP*$[f]$ *and MP models hold.*

Proof. If the PS model holds, then the QP$[f]$ model and the MP model hold. Assuming that both models hold, then we show that the PS model holds. Since $\{p_{ij}\}$ satisfy the QP$[f]$ model,

$$F\left(\frac{p_{ij}}{p_{ij}^{PS}}\right) = u_{1(i)} + u_{2(j)} + u_{12(ij)} \quad (i = 1, \ldots, r;\ j = 1, \ldots, r),$$

where $u_{12(ij)} = u_{12(i^*j^*)}$. Then we see

$$A_{ij} = u_{1(i)} - u_{1(i^*)} + u_{2(j)} - u_{2(j^*)}, \tag{2.3}$$

where

$$A_{ij} = F\left(\frac{p_{ij}}{p_{ij}^{PS}}\right) - F\left(\frac{p_{i^*j^*}}{p_{i^*j^*}^{PS}}\right).$$

The sum of right side of equation (2.3) multiplied by p_{ij} equals zero, because $\{p_{ij}\}$ satisfy the structure of MP model. Therefore,

$$\sum_{i=1}^{r}\sum_{j=1}^{r} p_{ij}A_{ij} = 0. \tag{2.4}$$

Since $A_{i^*j^*} = -A_{ij}$, the left side of equation (2.4) is

$$\begin{cases} \displaystyle\sum_{i=1}^{r/2}\sum_{j=1}^{r} \left(p_{ij} - p_{i^*j^*}\right) A_{ij} & (r : \text{even}), \\ \displaystyle\sum_{i=1}^{(r-1)/2}\sum_{j=1}^{r} \left(p_{ij} - p_{i^*j^*}\right) A_{ij} + \sum_{j=1}^{(r-1)/2} \left(p_{\frac{r+1}{2}j} - p_{\frac{r+1}{2}j^*}\right) A_{\frac{r+1}{2}j} & (r : \text{odd}). \end{cases}$$

Since the function F is a monotonically increasing function, $(p_{ij} - p_{i^*j^*})A_{ij}$ is greater than or equal to zero for any (i, j). Thus, we can obtain $\{p_{ij} = p_{i^*j^*}\}$. Namely, $\{p_{ij}\}$ satisfy the structure of point-symmetry. The proof is complete. □

§3. Case of three-way tables

3.1. Models

For an $r \times r \times r$ contingency table, let X_1, X_2, and X_3 denote the first, second, and third variables, respectively, and let $\Pr(X_1 = i, X_2 = j, X_3 = k) = p_{ijk}$. The point-symmetry ($\mathrm{PS}^3$) model can be expressed as

$$p_{ijk} = p_{i^*j^*k^*} \quad (1 \le i, j, k \le r),$$

where $i^* = r+1-i$ (Wall and Lienert, 1976). The PS^3 model may be expressed in a log-linear form

$$\begin{cases} \log p_{ijk} = u_{1(i)} + u_{2(j)} + u_{3(k)} \\ \qquad\qquad +u_{12(ij)} + u_{13(ik)} + u_{23(jk)} + u_{123(ijk)} \quad (1 \le i, j, k \le r), \\ \text{where for } d = 1, 2, 3 \text{ and } 1 \le s < t \le 3, \\ u_{d(i)} = u_{d(i^*)}, \quad u_{st(ij)} = u_{st(i^*j^*)}, \quad u_{123(ijk)} = u_{123(i^*j^*k^*)}. \end{cases}$$

See Tahata and Tomizawa (2008) for details.

The first-order marginal point-symmetry (MP^3_1) model is defined by

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

where $p_{i\cdot\cdot} = \sum_s \sum_t p_{ist}$, $p_{\cdot i\cdot} = \sum_s \sum_t p_{sit}$ and $p_{\cdot\cdot i} = \sum_s \sum_t p_{sti}$. This model indicates that the marginal distributions of X_k ($k = 1, 2, 3$) are point symmetric with respect to the midpoint of the categories.

The second-order marginal point-symmetry (MP^3_2) model is defined by

$$p_{ij\cdot} = p_{i^*j^*\cdot}, \quad p_{i\cdot j} = p_{i^*\cdot j^*} \quad \text{and} \quad p_{\cdot ij} = p_{\cdot i^*j^*} \quad (i = 1, \ldots, r;\ j = 1, \ldots, r),$$

where $p_{ij\cdot} = \sum_s p_{ijs}$, $p_{i\cdot j} = \sum_s p_{isj}$ and $p_{\cdot ij} = \sum_s p_{sij}$. This model indicates that the marginal distributions of X_s and X_t ($1 \le s < t \le 3$) are point symmetric with respect to the center point (when r is even) or the center cell (when r is odd) in the marginal $r \times r$ table.

Let f be a twice-differentiable and strictly convex function, and let $F(u) = df(u)/du$. Let $p^{PS}_{ijk} = (p_{ijk} + p_{i^*j^*k^*})/2$ for $1 \le i, j, k \le r$. We propose two new models, i.e., two kinds of quasi point-symmetry models based on f-divergence below.

First, we propose the first-order quasi point-symmetry ($\mathrm{QP}[f]^3_1$) model based on f-divergence as follows:

$$\begin{cases} p_{ijk} = p^{PS}_{ijk} F^{-1}\big(u_{1(i)} + u_{2(j)} + u_{3(k)} \\ \qquad\qquad +u_{12(ij)} + u_{13(ik)} + u_{23(jk)} + u_{123(ijk)}\big) \quad (1 \le i, j, k \le r), \\ \text{where for } 1 \le s < t \le 3, \\ u_{st(ij)} = u_{st(i^*j^*)}, \quad u_{123(ijk)} = u_{123(i^*j^*k^*)}. \end{cases}$$

The $\mathrm{QP}[f]_1^3$ model can also be expressed as

$$(3.1)\quad \begin{cases} \theta^{[f]}_{(i;j_1<j_2;k_1<k_2)} = \theta^{[f]}_{(i^*;j_2^*<j_1^*;k_2^*<k_1^*)} \\ \qquad\qquad (1 \le i \le r;\ 1 \le j_1 < j_2 \le r;\ 1 \le k_1 < k_2 \le r), \\ \theta^{[f]}_{(i_1<i_2;j;k_1<k_2)} = \theta^{[f]}_{(i_2^*<i_1^*;j^*;k_2^*<k_1^*)} \\ \qquad\qquad (1 \le i_1 < i_2 \le r;\ 1 \le j \le r;\ 1 \le k_1 < k_2 \le r), \\ \text{and} \\ \theta^{[f]}_{(i_1<i_2;j_1<j_2;k)} = \theta^{[f]}_{(i_2^*<i_1^*;j_2^*<j_1^*;k^*)} \\ \qquad\qquad (1 \le i_1 < i_2 \le r;\ 1 \le j_1 < j_2 \le r;\ 1 \le k \le r), \end{cases}$$

where

$$\theta^{[f]}_{(i;j_1<j_2;k_1<k_2)} = F\left(\frac{p_{ij_1k_1}}{p^{PS}_{ij_1k_1}}\right) + F\left(\frac{p_{ij_2k_2}}{p^{PS}_{ij_2k_2}}\right) - F\left(\frac{p_{ij_2k_1}}{p^{PS}_{ij_2k_1}}\right) - F\left(\frac{p_{ij_1k_2}}{p^{PS}_{ij_1k_2}}\right),$$

$$\theta^{[f]}_{(i_1<i_2;j;k_1<k_2)} = F\left(\frac{p_{i_1jk_1}}{p^{PS}_{i_1jk_1}}\right) + F\left(\frac{p_{i_2jk_2}}{p^{PS}_{i_2jk_2}}\right) - F\left(\frac{p_{i_2jk_1}}{p^{PS}_{i_2jk_1}}\right) - F\left(\frac{p_{i_1jk_2}}{p^{PS}_{i_1jk_2}}\right),$$

$$\theta^{[f]}_{(i_1<i_2;j_1<j_2;k)} = F\left(\frac{p_{i_1j_1k}}{p^{PS}_{i_1j_1k}}\right) + F\left(\frac{p_{i_2j_2k}}{p^{PS}_{i_2j_2k}}\right) - F\left(\frac{p_{i_2j_1k}}{p^{PS}_{i_2j_1k}}\right) - F\left(\frac{p_{i_1j_2k}}{p^{PS}_{i_1j_2k}}\right).$$

Tahata and Tomizawa (2008) proposed the first-order quasi point-symmetry (QP_1^3) model. This model has its characterization in terms of point-symmetry of odds ratios. Since equation (3.1) with $f(u) = u \log u,\ u > 0$ reduces to the structure of point-symmetry of odds ratios, we can see that the $\mathrm{QP}[f]_1^3$ model is equivalent to the QP_1^3 model when we set $f(u) = u \log u$. When we set $f(u) = (1-u)^2,\ u > 0$, we shall refer to the $\mathrm{QP}[f]_1^3$ model as the Pearsonian-QP_1^3 model.

Second, we propose the second-order quasi point-symmetry ($\mathrm{QP}[f]_2^3$) model based on f-divergence as follows:

$$\begin{cases} p_{ijk} = p^{PS}_{ijk} F^{-1}\left(u_{1(i)} + u_{2(j)} + u_{3(k)} \right. \\ \qquad\qquad \left. + u_{12(ij)} + u_{13(ik)} + u_{23(jk)} + u_{123(ijk)}\right) \quad (1 \le i, j, k \le r), \\ \text{where } u_{123(ijk)} = u_{123(i^*j^*k^*)}. \end{cases}$$

The $\mathrm{QP}[f]_2^3$ model can also be expressed as

$$(3.2)\quad \begin{cases} \theta^{[f]}_{(i_2;j_1<j_2;k_1<k_2)} - \theta^{[f]}_{(i_1;j_1<j_2;k_1<k_2)} = \\ \qquad\qquad \theta^{[f]}_{(i_2^*;j_2^*<j_1^*;k_2^*<k_1^*)} - \theta^{[f]}_{(i_1^*;j_2^*<j_1^*;k_2^*<k_1^*)}, \\ \text{or} \\ \theta^{[f]}_{(i_1<i_2;j_2;k_1<k_2)} - \theta^{[f]}_{(i_1<i_2;j_1;k_1<k_2)} = \\ \qquad\qquad \theta^{[f]}_{(i_2^*<i_1^*;j_2^*;k_2^*<k_1^*)} - \theta^{[f]}_{(i_2^*<i_1^*;j_1^*;k_2^*<k_1^*)}, \\ \text{or} \\ \theta^{[f]}_{(i_1<i_2;j_1<j_2;k_2)} - \theta^{[f]}_{(i_1<i_2;j_1<j_2;k_1)} = \\ \qquad\qquad \theta^{[f]}_{(i_2^*<i_1^*;j_2^*<j_1^*;k_2^*)} - \theta^{[f]}_{(i_2^*<i_1^*;j_2^*<j_1^*;k_1^*)}, \end{cases}$$

for $1 \le i_1 < i_2 \le r$; $1 \le j_1 < j_2 \le r$ and $1 \le k_1 < k_2 \le r$. Tahata and Tomizawa (2008) proposed the second-order quasi point-symmetry (QP_2^3) model. This model has its characterization in terms of point-symmetry of ratio of odds ratios. Since equation (3.2) with $f(u) = u \log u,\ u > 0$ reduces to the structure of point-symmetry of ratio of odds ratios, we can see that the $\mathrm{QP}[f]_2^3$ model is equivalent to the QP_2^3 model when we set $f(u) = u \log u$. When we set $f(u) = (1-u)^2,\ u > 0$, we shall refer to the $\mathrm{QP}[f]_2^3$ model as the Pearsonian-QP_2^3 model. We obtain following two theorems.

Theorem 3.1. *The $QP[f]_1^3$ model for $\{p_{ijk}\}$ minimizes the f-divergence between $\{p_{ijk}\}$ and $\left\{p_{ijk}^{PS}\right\}$ with the structure of the PS^3 model under the condition that $\{p_{i\cdot\cdot}\}$, $\{p_{\cdot j\cdot}\}$, $\{p_{\cdot\cdot k}\}$, $\{p_{ij\cdot} + p_{i^*j^*\cdot}\}$, $\{p_{i\cdot k} + p_{i^*\cdot k^*}\}$, $\{p_{\cdot jk} + p_{\cdot j^*k^*}\}$ and $\{p_{ijk} + p_{i^*j^*k^*}\}$ are given.*

Theorem 3.2. *The $QP[f]_2^3$ model for $\{p_{ijk}\}$ minimizes the f-divergence between $\{p_{ijk}\}$ and $\left\{p_{ijk}^{PS}\right\}$ with the structure of the PS^3 model under the condition that $\{p_{i\cdot\cdot}\}$, $\{p_{\cdot j\cdot}\}$, $\{p_{\cdot\cdot k}\}$, $\{p_{ij\cdot}\}$, $\{p_{i\cdot k}\}$, $\{p_{\cdot jk}\}$ and $\{p_{ijk} + p_{i^*j^*k^*}\}$ are given.*

The proofs of Theorems 3.1 and 3.2 are omitted because it can be obtained in a similar way as the proof of Theorem 2.1.

3.2. Decomposition of point-symmetry

Tahata and Tomizawa (2008) gave the decomposition of the PS^3 model for the $r \times r \times r$ table. We obtain the following theorem, which includes the result of

Tahata and Tomizawa (2008).

Theorem 3.3. *For an $r \times r \times r$ table and h fixed $(h = 1, 2)$, the PS^3 model holds if and only if both the $QP[f]_h^3$ and MP_h^3 models hold.*

Proof. We give the proof when r is odd and $h = 2$. If the PS^3 model holds, then the QP$[f]_2^3$ model and the MP$_2^3$ model hold. Assuming that both the QP$[f]_2^3$ and MP$_2^3$ models hold, then we show that the PS^3 model holds. Since $\{p_{ijk}\}$ satisfy the QP$[f]_2^3$ model, we see

$$\begin{aligned} A_{ijk} &= u_{1(i)} - u_{1(i^*)} + u_{2(j)} - u_{2(j^*)} + u_{3(k)} - u_{3(k^*)} \\ &\quad + u_{12(ij)} - u_{12(i^*j^*)} + u_{13(ik)} - u_{13(i^*k^*)} + u_{23(jk)} - u_{23(j^*k^*)}, \end{aligned} \tag{3.3}$$

for $1 \leq i, j, k \leq r$, where

$$A_{ijk} = F\left(\frac{p_{ijk}}{p_{ijk}^{PS}}\right) - F\left(\frac{p_{i^*j^*k^*}}{p_{i^*j^*k^*}^{PS}}\right).$$

The sum of right side of equation (3.3) multiplied by p_{ijk} equals zero, because $\{p_{ijk}\}$ satisfy the structure of the MP$_2^3$ model. Therefore,

$$\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r} p_{ijk} A_{ijk} = 0. \tag{3.4}$$

The left side of equation (3.4) is

$$\begin{aligned} &\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{(r-1)/2} \left(p_{ijk} - p_{i^*j^*k^*}\right) A_{ijk} \\ &\qquad + \sum_{i=1}^{(r-1)/2}\sum_{j=1}^{r} \left(p_{ij,\frac{(r+1)}{2}} - p_{i^*j^*,\frac{(r+1)}{2}}\right) A_{ij,\frac{(r+1)}{2}} \\ &\qquad\quad + \sum_{j=1}^{(r-1)/2} \left(p_{\frac{(r+1)}{2},j,\frac{(r+1)}{2}} - p_{\frac{(r+1)}{2},j^*,\frac{(r+1)}{2}}\right) A_{\frac{(r+1)}{2},j,\frac{(r+1)}{2}}. \end{aligned}$$

Since the function F is a monotonically increasing function, $(p_{ijk} - p_{i^*j^*k^*})A_{ijk}$ is greater than or equal to zero for any (i, j, k). Thus we obtain $\{p_{ijk} = p_{i^*j^*k^*}\}$. Namely, $\{p_{ijk}\}$ satisfy the structure of point-symmetry. The other cases can be proved in a similar way. The proof is complete. □

§4. Case of multi-way tables

4.1. Models

Consider an r^T contingency table ($T \geq 2$), let $p_{\boldsymbol{i}}$ denote the probability that an observation will fall in the $\boldsymbol{i}$th cell of the table, where $\boldsymbol{i} = (i_1, \ldots, i_T)$ for $i_k = 1, \ldots, r$ ($k = 1, \ldots, T$).

The point-symmetry (PS^T) model is defined by

$$p_{\boldsymbol{i}} = p_{\boldsymbol{i}^*} \quad \text{for any } \boldsymbol{i},$$

where $\boldsymbol{i}^* = (i_1^*, \ldots, i_T^*)$ for $i_k^* = r + 1 - i_k$ ($k = 1, \ldots, T$) (Wall and Lienert, 1976). The PS^T model can be expressed in a log-linear form

$$\begin{cases} \log p_{\boldsymbol{i}} = \displaystyle\sum_{k=1}^{T} u_{k(i_k)} + \sum\sum_{1 \leq k_1 < k_2 \leq T} u_{k_1 k_2 (i_{k_1}, i_{k_2})} \\ \qquad + \cdots + \displaystyle\sum \cdots \sum_{1 \leq k_1 < \cdots < k_{T-1} \leq T} u_{k_1 \ldots k_{T-1}(i_{k_1}, \ldots, i_{k_{T-1}})} + u_{12\ldots T(\boldsymbol{i})} \text{ for any } \boldsymbol{i}, \\ \text{where } u_{k_1 k_2 \ldots k_l (i_{k_1}, i_{k_2}, \ldots, i_{k_l})} = u_{k_1 k_2 \ldots k_l (i_{k_1}^*, i_{k_2}^*, \ldots, i_{k_l}^*)} \\ \qquad (l = 1, \ldots, T; 1 \leq k_1 < k_2 < \cdots < k_l \leq T). \end{cases}$$

See Tahata and Tomizawa (2008) for details.

Denote the hth-order marginal probability $\Pr(X_{s_1} = i_1, \ldots, X_{s_h} = i_h)$ by $p_{\boldsymbol{i}_h}^{\boldsymbol{s}_h}$, where $\boldsymbol{s}_h = (s_1, \ldots, s_h)$ and $\boldsymbol{i}_h = (i_1, \ldots, i_h)$ with $1 \leq s_1 < \cdots < s_h \leq T$ and $i_k = 1, \ldots, r$ ($k = 1, \ldots, h$) for h ($h = 1, \ldots, T-1$). For a fixed h ($h = 1, \ldots, T-1$), the hth-order marginal point-symmetry (MP_h^T) model is defined by

$$p_{\boldsymbol{i}_h}^{\boldsymbol{s}_h} = p_{\boldsymbol{i}_h^*}^{\boldsymbol{s}_h} \quad \text{for any } \boldsymbol{s}_h = (s_1, \ldots, s_h),$$

where $\boldsymbol{i}_h = (i_1, \ldots, i_h)$ and $\boldsymbol{i}_h^* = (i_1^*, \ldots, i_h^*)$ (Tahata and Tomizawa, 2008).

Let f be a twice-differentiable and strictly convex function, and let $F(u) = df(u)/du$. Let $p_{\boldsymbol{i}}^{PS} = (p_{\boldsymbol{i}} + p_{\boldsymbol{i}^*})/2$ for any $\boldsymbol{i}$. For a fixed h ($h = 1, \ldots, T-1$), we propose the hth-order quasi point-symmetry ($\mathrm{QP}[f]_h^T$) model based on f-divergence as follows:

$$\begin{cases} p_{\boldsymbol{i}} = p_{\boldsymbol{i}}^{PS} F^{-1} \left(\displaystyle\sum_{k=1}^{T} u_{k(i_k)} + \sum\sum_{1 \leq k_1 < k_2 \leq T} u_{k_1 k_2 (i_{k_1}, i_{k_2})} \right. \\ \qquad \left. + \cdots + \displaystyle\sum \cdots \sum_{1 \leq k_1 < \cdots < k_{T-1} \leq T} u_{k_1 \ldots k_{T-1}(i_{k_1}, \ldots, i_{k_{T-1}})} + u_{12\ldots T(\boldsymbol{i})} \right) \text{ for any } \boldsymbol{i}, \\ \text{where } u_{k_1 k_2 \ldots k_l (i_{k_1}, i_{k_2}, \ldots, i_{k_l})} = u_{k_1 k_2 \ldots k_l (i_{k_1}^*, i_{k_2}^*, \ldots, i_{k_l}^*)} \\ \qquad (l = h+1, \ldots, T; 1 \leq k_1 < k_2 < \cdots < k_l \leq T). \end{cases}$$

When we set $f(u) = u \log u,\ u > 0$, the $\mathrm{QP}[f]_h^T$ model is equivalent to the hth-order quasi point-symmetry model proposed by Tahata and Tomizawa (2008). When we set $f(u) = (1-u)^2,\ u > 0$, we shall refer to the $\mathrm{QP}[f]_h^T$ model as the Pearsonian-QP_h^T model. We obtain the following theorem.

Theorem 4.1. *For an r^T table and h fixed ($h = 1, \ldots, T-1$), the $QP[f]_h^T$ model for $\{p_{\boldsymbol{i}}\}$ minimizes the f-divergence between $\{p_{\boldsymbol{i}}\}$ and $\left\{p_{\boldsymbol{i}}^{PS}\right\}$ with the structure of the PS^T model under the condition that $\left\{p_{\boldsymbol{i}_k}^{\boldsymbol{s}_k}\right\}$ for any $\boldsymbol{s}_k = (s_1, \ldots, s_k)$ and $\boldsymbol{i}_k = (i_1, \ldots, i_k)$, $k = 1, \ldots, h$, are given, and $\left\{p_{\boldsymbol{i}_l}^{\boldsymbol{s}_l} + p_{\boldsymbol{i}_l^*}^{\boldsymbol{s}_l}\right\}$ for any $\boldsymbol{s}_l = (s_1, \ldots, s_l)$, $\boldsymbol{i}_l = (i_1, \ldots, i_l)$ and $\boldsymbol{i}_l^* = (i_1^*, \ldots, i_l^*)$, $l = h+1, \ldots, T$, are given.*

The proof of Theorem 4.1 is omitted because it can be obtained in a similar way as the proof of Theorem 2.1.

Tahata and Tomizawa (2008) showed the decomposition of the PS^T model for an r^T table. We obtain the following theorem, which includes the result of Tahata and Tomizawa (2008).

Theorem 4.2. *For an r^T table and h fixed ($h = 1, \ldots, T-1$), the PS^T model holds if and only if both the $QP[f]_h^T$ and MP_h^T models hold.*

The proof of Theorem 4.2 is omitted because it can be obtained in a similar way as the proof of Theorem 3.3.

§5. Properties of test statistics

Let $n_{i_1 \ldots i_T}$ denote the observed frequency in the $(i_1, \ldots, i_T)$th cell of the r^T table. Assume that a multinomial distribution is applied to the r^T table. The maximum likelihood estimates (MLEs) of expected frequencies under each model could be obtained by using the Newton-Raphson method in the log-likelihood equation. Each model can be tested for goodness-of-fit by, e.g., the likelihood ratio chi-squared statistic (denoted by G^2) with corresponding degrees of freedom (df). The number of df for the considered models are given in Table 1.

The test statistic G^2 for model H is given by

$$G^2(H) = 2 \sum_{i_1=1}^{r} \cdots \sum_{i_T=1}^{r} n_{i_1 \ldots i_T} \log \left(\frac{n_{i_1 \ldots i_T}}{\hat{m}_{i_1 \ldots i_T}} \right),$$

where $\hat{m}_{i_1 \ldots i_T}$ is the MLE of expected frequency $m_{i_1 \ldots i_T}$ under model H. Consider two nested models, say H_1 and H_2, such that model H_1 is a special case of model H_2, if model H_1 holds, then model H_2 also holds. Let v_1 and v_2 denote the df for models H_1 and H_2, respectively. For testing that model H_1 holds assuming that model H_2 holds true, the likelihood ratio statistics is given as $G^2(H_1 \mid H_2) = G^2(H_1) - G^2(H_2)$. Under the null hypothesis this test statistic has an asymptotic chi-square distribution with $v_1 - v_2$ df.

Aitchison (1962) discussed the asymptotic separability of models. Also the similar property of models is described by Darroch and Silvey (1963) and Read (1977). (See also, Tahata and Tomizawa, 2008; Tomizawa, 1993; Tomizawa and Tahata, 2007). Generally suppose that model H_3 holds if and only if both model H_1 and model H_2 hold. When the test statistic for goodness-of-fit of model H_3 is asymptotically equivalent to the sum of those for model H_1 and model H_2, if both model H_1 and model H_2 are accepted (at the α significance level) with high probability, then model H_3 would be accepted. However, when it does not hold, it is quite possible for an incompatible situation to arise where both model H_1 and model H_2 are accepted but model H_3 is rejected with high probability. Thus, we consider the partitions of test statistics.

First, we can obtain the following theorem. Note that the theorem with $f(u) = u \log u$ is given by Tahata and Tomizawa (2008).

Theorem 5.1. *For an $r \times r$ table, $G^2(PS)$ is asymptotically equivalent to the sum of $G^2(QP[f])$ and $G^2(MP)$ under the PS model.*

Proof. Consider the case that r is odd. The QP[f] model may be expressed as

$$p_{ij} = p_{ij}^{PS} F^{-1}(u + u'_{1(i)} + u'_{2(j)} + u'_{12(ij)}) \quad (i = 1, \ldots, r;\ j = 1, \ldots, r),$$

where $u'_{12(ij)} = u'_{12(i^*j^*)}$. Without loss of generality, we may set $\sum_i u'_{1(i)} = \sum_j u'_{2(j)} = \sum_i u'_{12(ij)} = \sum_j u'_{12(ij)} = 0$. Let

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

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

$$\boldsymbol{\beta} = (u, u'_{1(1)}, \ldots, u'_{1(r-1)}, u'_{2(1)}, \ldots, u'_{2(r-1)}, \boldsymbol{\beta}_{12})^t,$$

where "t" denotes the transpose, and $\boldsymbol{\beta}_{12}$ is the $1 \times (r-1)^2/2$ vector of $u'_{12(ij)}$ for $i = 1, \ldots, (r-1)/2$ and $j = 1, \ldots, r-1$.

Then the QP[f] model is also expressed as

$$\boldsymbol{F}\left(\frac{\boldsymbol{p}}{\boldsymbol{p}^{PS}}\right) = \boldsymbol{X}\boldsymbol{\beta} = (\boldsymbol{1}_{r^2}, \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_{12})\boldsymbol{\beta},$$

where $\boldsymbol{F}\left(\boldsymbol{p}/\boldsymbol{p}^{PS}\right)$ is the $r^2 \times 1$ vector with components $F\left(p_{ij}/p_{ij}^{PS}\right)$, $\boldsymbol{X}$ is the $r^2 \times K$ matrix with $K = (r^2+2r-1)/2$ and $\boldsymbol{1}_s$ is the $s \times 1$ vector of 1 elements,

$$\boldsymbol{X}_1 = \begin{pmatrix} \boldsymbol{I}_{r-1} \otimes \boldsymbol{1}_r \\ -\boldsymbol{1}_r \boldsymbol{1}_{r-1}^t \end{pmatrix}; \quad r^2 \times (r-1) \text{ matrix},$$

$$\boldsymbol{X}_2 = \boldsymbol{1}_r \otimes \begin{pmatrix} \boldsymbol{I}_{r-1} \\ -\boldsymbol{1}_{r-1}^t \end{pmatrix}; \quad r^2 \times (r-1) \text{ matrix},$$

and $\boldsymbol{X}_{12}$ is the $r^2 \times (r-1)^2/2$ matrix determined from the structure of the QP[f] model, $\boldsymbol{I}_s$ is the $s \times s$ identity matrix, and "$\otimes$" denotes the Kronecker product. Note that the matrix $\boldsymbol{X}$ has full column rank which is K.

We denote the linear space spanned by the columns of the matrix $\boldsymbol{X}$ by $S(\boldsymbol{X})$ with dimension K. Let $\boldsymbol{U}$ be an $r^2 \times d_1$ full column rank matrix, where $d_1 = r^2 - K = (r-1)^2/2$, such that the linear space spanned by the columns of $\boldsymbol{U}$, i.e., $S(\boldsymbol{U})$, is orthogonal complement of the space $S(\boldsymbol{X})$. Thus, $\boldsymbol{U}^t\boldsymbol{X} = \boldsymbol{O}_{d_1,K}$, where $\boldsymbol{O}_{d_1,K}$ is the $d_1 \times K$ zero matrix. Therefore the QP[f] model is expressed as

$$\boldsymbol{h}_1(\boldsymbol{p}) = \boldsymbol{0}_{d_1},$$

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

$$\boldsymbol{h}_1(\boldsymbol{p}) = \boldsymbol{U}^t\boldsymbol{F}\left(\frac{\boldsymbol{p}}{\boldsymbol{p}^{PS}}\right).$$

The MP model may be expressed as

$$\boldsymbol{h}_2(\boldsymbol{p}) = \boldsymbol{0}_{d_2},$$

where $d_2 = r - 1$,

$$\boldsymbol{h}_2(\boldsymbol{p}) = \boldsymbol{M}\boldsymbol{p} = \begin{pmatrix} \boldsymbol{M}_1 \\ \boldsymbol{M}_2 \end{pmatrix} \boldsymbol{p}$$

with $\boldsymbol{M}$ being the $d_2 \times r^2$ matrix and for $k = 1, 2$,

$$\boldsymbol{M}_k = \begin{pmatrix} \boldsymbol{x}_{k(1)}^t \\ \boldsymbol{x}_{k(2)}^t - \boldsymbol{x}_{k(2^*)}^t \\ \vdots \\ \boldsymbol{x}_{k(\frac{r-1}{2})}^t - \boldsymbol{x}_{k((\frac{r-1}{2})^*)}^t \end{pmatrix}; \quad (r-1)/2 \times r^2 \text{ matrix},$$

where $\boldsymbol{x}_{k(i)}$ is the $r^2 \times 1$ vector which is the ith column vector of $\boldsymbol{X}_k$ ($i = 1, \ldots, r-1$). Thus the column vectors of $\boldsymbol{M}^t$ belong to the space $S(\boldsymbol{X})$, i.e., $S(\boldsymbol{M}^t) \subset S(\boldsymbol{X})$. Hence $\boldsymbol{M}\boldsymbol{U} = \boldsymbol{O}_{d_2 d_1}$. From Theorem 2.2, the PS model may be expressed as

$$\boldsymbol{h}_3(\boldsymbol{p}) = \boldsymbol{0}_{d_3},$$

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

$$\boldsymbol{h}_3(\boldsymbol{p}) = (\boldsymbol{h}_1(\boldsymbol{p})^t, \boldsymbol{h}_2(\boldsymbol{p})^t)^t.$$

Note that $\boldsymbol{h}_s(\boldsymbol{p})$, $s = 1, 2, 3$, are the vectors of order $d_s \times 1$, and d_s, $s = 1, 2, 3$, are the numbers of df for testing goodness-of-fit of the QP[f], MP and PS models, respectively.

Let $\boldsymbol{H}_s(\boldsymbol{p})$, $s = 1, 2, 3$, denote the $d_s \times r^2$ matrix of partial derivatives of $\boldsymbol{h}_s(\boldsymbol{p})$ with respect to $\boldsymbol{p}$, i.e., $\boldsymbol{H}_s(\boldsymbol{p}) = \partial \boldsymbol{h}_s(\boldsymbol{p})/\partial \boldsymbol{p}^t$. Let $\boldsymbol{\Sigma}(\boldsymbol{p}) = diag(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^t$, where $diag(\boldsymbol{p})$ denotes a diagonal matrix with ith component of $\boldsymbol{p}$ as ith diagonal component. Let $\hat{\boldsymbol{p}}$ denote $\boldsymbol{p}$ with p_{ij} replaced by $\hat{p}_{ij}$, where $\hat{p}_{ij} = n_{ij}/n$ with $n = \sum\sum n_{ij}$. Then $\sqrt{n}(\hat{\boldsymbol{p}} - \boldsymbol{p})$ has asymptotically a normal distribution with mean vector $\boldsymbol{0}_{r^2}$ and covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{p})$. Using the delta method, $\sqrt{n}(\boldsymbol{h}_3(\hat{\boldsymbol{p}}) - \boldsymbol{h}_3(\boldsymbol{p}))$ has asymptotically a normal distribution with mean $\boldsymbol{0}_{d_3}$ and covariance matrix

$$\boldsymbol{H}_3(\boldsymbol{p})\boldsymbol{\Sigma}(\boldsymbol{p})\boldsymbol{H}_3^t(\boldsymbol{p}) = \begin{pmatrix} \boldsymbol{H}_1(\boldsymbol{p})\boldsymbol{\Sigma}(\boldsymbol{p})\boldsymbol{H}_1^t(\boldsymbol{p}) & \boldsymbol{H}_1(\boldsymbol{p})\boldsymbol{\Sigma}(\boldsymbol{p})\boldsymbol{H}_2^t(\boldsymbol{p}) \\ \boldsymbol{H}_2(\boldsymbol{p})\boldsymbol{\Sigma}(\boldsymbol{p})\boldsymbol{H}_1^t(\boldsymbol{p}) & \boldsymbol{H}_2(\boldsymbol{p})\boldsymbol{\Sigma}(\boldsymbol{p})\boldsymbol{H}_2^t(\boldsymbol{p}) \end{pmatrix}.$$

We obtain

$$\boldsymbol{H}_1(\boldsymbol{p}) = \boldsymbol{U}^t \left(diag(\boldsymbol{a}) - \boldsymbol{J} diag(\boldsymbol{b})\right),$$

where $\boldsymbol{J}$ is the $r^2 \times r^2$ matrix with 1 in the $(i, r^2 + 1 - i)$th element and 0 otherwise, and

$$\begin{aligned} \boldsymbol{a} &= (a_{11}, \ldots, a_{1r}, a_{21}, \ldots, a_{2r}, \ldots, a_{r1}, \ldots, a_{rr})^t, \\ \boldsymbol{b} &= (b_{11}, \ldots, b_{1r}, b_{21}, \ldots, b_{2r}, \ldots, b_{r1}, \ldots, b_{rr})^t, \\ a_{ij} &= \frac{p_{i^*j^*}}{2\left(p_{ij}^{PS}\right)^2} f''\left(\frac{p_{ij}}{p_{ij}^{PS}}\right), \\ b_{ij} &= \frac{p_{i^*j^*}}{2\left(p_{ij}^{PS}\right)^2} f''\left(\frac{p_{i^*j^*}}{p_{ij}^{PS}}\right), \end{aligned}$$

with $f''(u) = dF(u)/du$. Under the PS model, we see

$$\begin{aligned} \boldsymbol{H}_1(\boldsymbol{p})\boldsymbol{p} &= \boldsymbol{0}_{d_1}, \\ \boldsymbol{H}_1(\boldsymbol{p}) diag(\boldsymbol{p}) &= c\boldsymbol{U}^t \left(\boldsymbol{I} - \boldsymbol{J}\right), \end{aligned}$$

with $c = f''(1)/2$. Also we see

$$\boldsymbol{H}_2(\boldsymbol{p}) = \boldsymbol{M}.$$

Noting that $\boldsymbol{J}\boldsymbol{M}^t = -\boldsymbol{M}^t$ and $\boldsymbol{M}\boldsymbol{U} = \boldsymbol{O}_{d_2 d_1}$, we obtain that under the PS model

$$\boldsymbol{H}_1(\boldsymbol{p})\boldsymbol{\Sigma}(\boldsymbol{p})\boldsymbol{H}_2^t(\boldsymbol{p}) = \boldsymbol{O}_{d_1 d_2}.$$

Thus, under the PS model, we obtain $W_3 = W_1 + W_2$, where

$$W_s = n\boldsymbol{h}_s(\boldsymbol{p})^t(\boldsymbol{H}_s(\boldsymbol{p})\boldsymbol{\Sigma}(\boldsymbol{p})\boldsymbol{H}_s^t(\boldsymbol{p}))^{-1}\boldsymbol{h}_s(\boldsymbol{p}) \quad (s = 1, 2, 3).$$

Wald statistic $\hat{W}_s$, i.e., W_s with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$, has asymptotically a chi-squared distribution with d_s df under the corresponding model. From the asymptotic equivalence of the Wald statistic and likelihood ratio statistic (Rao, 1973, Sec. 6e. 3), we obtain Theorem 5.1 when r is odd. In an analogous way we obtain Theorem 5.1 when r is even. The proof is complete. □

Second, we shall consider the case of $T = 3$, i.e., $r \times r \times r$ table. We can obtain the following theorem.

Theorem 5.2. *For an $r \times r \times r$ table and a fixed h ($h = 1, 2$), $G^2(PS^3)$ is asymptotically equivalent to the sum of $G^2(QP[f]_h^3)$ and $G^2(MP_h^3)$ under the PS^3 model.*

Proof. We shall consider the proof when r is odd and $h = 2$. The QP$[f]_2^3$ model may be expressed as

$$p_{ijk} = p_{ijk}^{PS} F^{-1}\Big(u + u'_{1(i)} + u'_{2(j)} + u'_{3(k)} + u'_{12(ij)} + u'_{13(ik)} + u'_{23(jk)} + u'_{123(ijk)}\Big) \quad (1 \leq i, j, k \leq r),$$

where $u'_{123(ijk)} = u'_{123(i^*j^*k^*)}$. In the QP$[f]_2^3$ model, without loss of generality, we may set $\sum_i u'_{m(i)} = 0$ $(m = 1, 2, 3)$, $\sum_i u'_{st(ij)} = \sum_j u'_{st(ij)} = 0$ $(1 \leq s < t \leq 3)$, and $\sum_i u'_{123(ijk)} = \sum_j u'_{123(ijk)} = \sum_k u'_{123(ijk)} = 0$. Let

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

$$\boldsymbol{p}^{PS} = (p_{111}^{PS}, \ldots, p_{1r1}^{PS}, \ldots, p_{r11}^{PS}, \ldots, p_{rr1}^{PS}, p_{112}^{PS}, \ldots, p_{1r2}^{PS}, \ldots, p_{r12}^{PS}, \ldots, p_{rr2}^{PS}, \ldots, p_{11r}^{PS}, \ldots, p_{1rr}^{PS}, \ldots, p_{r1r}^{PS}, \ldots, p_{rrr}^{PS})^t,$$

$$\boldsymbol{\beta} = (u, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\beta}_{12}, \boldsymbol{\beta}_{13}, \boldsymbol{\beta}_{23}, \boldsymbol{\beta}_{123})^t,$$

where

$$\boldsymbol{\beta}_m = (u'_{m(1)}, \ldots, u'_{m(r-1)}) \quad (m = 1, 2, 3),$$

$$\boldsymbol{\beta}_{st} = (u'_{st(11)}, \ldots, u'_{st(1,r-1)}, u'_{st(21)}, \ldots, u'_{st(2,r-1)}, \ldots, u'_{st(r-1,1)}, \ldots, u'_{st(r-1,r-1)}) \quad (1 \leq s < t \leq 3),$$

and $\boldsymbol{\beta}_{123}$ is the $(r-1)^3/2 \times 1$ vector of $u'_{123(ijk)}$. Then the QP$[f]_2^3$ model is expressed as

$$\boldsymbol{F}\left(\frac{\boldsymbol{p}}{\boldsymbol{p}^{PS}}\right) = \boldsymbol{X}\boldsymbol{\beta} = (\mathbf{1}_{r^3}, \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3, \boldsymbol{X}_{12}, \boldsymbol{X}_{13}, \boldsymbol{X}_{23}, \boldsymbol{X}_{123})\boldsymbol{\beta},$$

where $\boldsymbol{F}\left(\boldsymbol{p}/\boldsymbol{p}^{PS}\right)$ is the $r^3 \times 1$ vector with components $F\left(p_{ijk}/p_{ijk}^{PS}\right)$, where $\boldsymbol{X}$ is the $r^3 \times K$ matrix with $K = (r^3 + 3r^2 - 3r + 1)/2$,

$$\boldsymbol{X}_1 = \mathbf{1}_r \otimes \begin{pmatrix} \boldsymbol{I}_{r-1} \otimes \mathbf{1}_r \\ -\mathbf{1}_r \mathbf{1}_{r-1}^t \end{pmatrix}; \quad r^3 \times (r-1) \text{ matrix},$$

$$\boldsymbol{X}_2 = \mathbf{1}_{r^2} \otimes \begin{pmatrix} \boldsymbol{I}_{r-1} \\ -\mathbf{1}_{r-1}^t \end{pmatrix}; \quad r^3 \times (r-1) \text{ matrix},$$

$$\boldsymbol{X}_3 = \begin{pmatrix} \boldsymbol{I}_{r-1} \otimes \mathbf{1}_{r^2} \\ -\mathbf{1}_{r^2} \mathbf{1}_{r-1}^t \end{pmatrix}; \quad r^3 \times (r-1) \text{ matrix},$$

$$\boldsymbol{X}_{12} = \mathbf{1}_r \otimes \begin{pmatrix} \boldsymbol{I}_{r-1} \otimes \begin{pmatrix} \boldsymbol{I}_{r-1} \\ -\mathbf{1}_{r-1}^t \end{pmatrix} \\ \mathbf{1}_{r-1}^t \otimes \begin{pmatrix} -\boldsymbol{I}_{r-1} \\ \mathbf{1}_{r-1}^t \end{pmatrix} \end{pmatrix}; \quad r^3 \times (r-1)^2 \text{ matrix},$$

$$\boldsymbol{X}_{13} = \begin{pmatrix} \boldsymbol{A}_1 \\ \vdots \\ \boldsymbol{A}_{r-1} \\ \begin{pmatrix} \boldsymbol{I}_{r-1} \otimes \left(-\mathbf{1}_r \mathbf{1}_{r-1}^t\right) \\ \mathbf{1}_r \mathbf{1}_{(r-1)^2}^t \end{pmatrix} \end{pmatrix}; \quad r^3 \times (r-1)^2 \text{ matrix},$$

where

$$\boldsymbol{A}_i = \begin{pmatrix} \boldsymbol{I}_{r-1} \otimes \boldsymbol{C}_i \\ -\mathbf{1}_{r-1}^t \otimes \boldsymbol{C}_i \end{pmatrix}; \quad r^2 \times (r-1)^2 \text{ matrix},$$

$$\boldsymbol{C}_i = \begin{pmatrix} \boldsymbol{O}_{r,i-1} & \mathbf{1}_r & \boldsymbol{O}_{r,r-1-i} \end{pmatrix}; \quad r \times (r-1) \text{ matrix},$$

with

$$\boldsymbol{C}_1 = \begin{pmatrix} \mathbf{1}_r & \boldsymbol{O}_{r,r-2} \end{pmatrix}, \quad \boldsymbol{C}_{r-1} = \begin{pmatrix} \boldsymbol{O}_{r,r-2} & \mathbf{1}_r \end{pmatrix},$$

and

$$\boldsymbol{X}_{23} = \begin{pmatrix} \boldsymbol{D}_{11} & \cdots & \boldsymbol{D}_{r-1,1} \\ & \cdots & \\ \boldsymbol{D}_{1,r-1} & \cdots & \boldsymbol{D}_{r-1,r-1} \\ \multicolumn{3}{c}{\mathbf{1}_r \otimes \begin{pmatrix} \boldsymbol{I}_{r-1} \otimes -\mathbf{1}_{r-1}^t \\ \mathbf{1}_{(r-1)^2}^t \end{pmatrix}} \end{pmatrix}; \quad r^3 \times (r-1)^2 \text{ matrix},$$

where

$$\boldsymbol{D}_{ij} = \boldsymbol{1}_r \otimes \begin{pmatrix} \boldsymbol{E}_{ij} \\ \boldsymbol{G}_j \end{pmatrix}; \quad r^2 \times (r-1) \text{ matrix},$$

$$\boldsymbol{G}_j = \begin{pmatrix} \boldsymbol{0}^t_{j-1} & -1 & \boldsymbol{0}^t_{r-1-j} \end{pmatrix}; \quad 1 \times (r-1) \text{ vector},$$

with

$$\boldsymbol{G}_1 = \begin{pmatrix} -1 & \boldsymbol{0}^t_{r-2} \end{pmatrix}, \quad \boldsymbol{G}_{r-1} = \begin{pmatrix} \boldsymbol{0}^t_{r-2} & -1 \end{pmatrix},$$

$\boldsymbol{E}_{ij}$ is the $(r-1) \times (r-1)$ matrix with 1 in the (i,j)th element, and $\boldsymbol{X}_{123}$ is the $r^3 \times (r-1)^3/2$ matrix determined from the structure of the $\mathrm{QP}[f]^3_2$ model. Note that the matrix $\boldsymbol{X}$ has full column rank which is K.

We denote the linear space spanned by the columns of the matrix $\boldsymbol{X}$ by $S(\boldsymbol{X})$ with dimension K. Let $\boldsymbol{U}$ be an $r^3 \times d_1$ full column rank matrix, where $d_1 = r^3 - K = (r-1)^3/2$, such that the linear space spanned by the columns of $\boldsymbol{U}$, i.e., $S(\boldsymbol{U})$, is orthogonal complement of the space $S(\boldsymbol{X})$. Thus, $\boldsymbol{U}^t\boldsymbol{X} = \boldsymbol{O}_{d_1,K}$. Therefore the $\mathrm{QP}[f]^3_2$ model is expressed as

$$\boldsymbol{h}_1(\boldsymbol{p}) = \boldsymbol{U}^t\boldsymbol{F}\left(\frac{\boldsymbol{p}}{\boldsymbol{p}^{PS}}\right) = \boldsymbol{0}_{d_1}.$$

The MP^3_2 model may be expressed as

$$\boldsymbol{h}_2(\boldsymbol{p}) = \boldsymbol{0}_{d_2},$$

where $d_2 = 3r(r-1)/2$,

$$\boldsymbol{h}_2(\boldsymbol{p}) = \boldsymbol{M}\boldsymbol{p} = \begin{pmatrix} \boldsymbol{M}_1 \\ \boldsymbol{M}_2 \\ \boldsymbol{M}_3 \\ \boldsymbol{M}_{12} \\ \boldsymbol{M}_{13} \\ \boldsymbol{M}_{23} \end{pmatrix} \boldsymbol{p},$$

with $\boldsymbol{M}$ being the $d_2 \times r^3$ matrix and for $k = 1,2,3$,

$$\boldsymbol{M}_k = \begin{pmatrix} \boldsymbol{x}^t_{k(1)} \\ \boldsymbol{x}^t_{k(2)} - \boldsymbol{x}^t_{k(2^*)} \\ \vdots \\ \boldsymbol{x}^t_{k(\frac{r-1}{2})} - \boldsymbol{x}^t_{k((\frac{r-1}{2})^*)} \end{pmatrix}; \quad (r-1)/2 \times r^3 \text{ matrix},$$

and for $1 \le k < l \le 3$,

$$\boldsymbol{M}_{kl} = \begin{pmatrix} \boldsymbol{a}^t_{kl(11)} - \boldsymbol{a}^t_{kl(1^*1^*)} \\ \vdots \\ \boldsymbol{a}^t_{kl(1,r-1)} - \boldsymbol{a}^t_{kl(1^*,(r-1)^*)} \\ \vdots \\ \boldsymbol{a}^t_{kl(\frac{r-1}{2},r-1)} - \boldsymbol{a}^t_{kl((\frac{r-1}{2})^*,(r-1)^*)} \end{pmatrix}; \quad (r-1)^2/2 \times r^3 \text{ matrix},$$

where

$$\boldsymbol{a}_{kl(ij)} = \frac{1}{r}\boldsymbol{x}_{k(i)} + \frac{1}{r}\boldsymbol{x}_{l(j)} + \boldsymbol{x}_{kl(ij)} - \frac{1}{r}\sum_{m=1}^{r-1}\left(\boldsymbol{x}_{kl(im)} + \boldsymbol{x}_{kl(mj)}\right),$$

and we set for $i = 1, \ldots, r$,

$$\boldsymbol{x}_{k(r)} = \boldsymbol{x}_{l(r)} = \boldsymbol{x}_{kl(ir)} = \boldsymbol{x}_{kl(ri)} = \mathbf{0}_{r^3}.$$

Note that $\boldsymbol{x}_{k(i)}$ is the $r^3 \times 1$ column vector in $\boldsymbol{X}_k$ shouldering $u_{k(i)}$ ($k = 1, 2, 3;\ i = 1, \ldots, r-1$) and $\boldsymbol{x}_{kl(ij)}$ is the $r^3 \times 1$ vector in $\boldsymbol{X}_{kl}$ shouldering $u_{kl(ij)}$ ($1 \le k < l \le 3;\ 1 \le i, j \le r-1$). Thus the column vectors of $\boldsymbol{M}^t$ belong to the space $S(\boldsymbol{X})$, i.e., $S(\boldsymbol{M}^t) \subset S(\boldsymbol{X})$. Hence $\boldsymbol{M}\boldsymbol{U} = \boldsymbol{O}_{d_2 d_1}$. From Theorem 3.3, the PS^3 model may be expressed as

$$\boldsymbol{h}_3(\boldsymbol{p}) = \mathbf{0}_{d_3},$$

where $d_3 = d_1 + d_2 = (r^3 - 1)/2$,

$$\boldsymbol{h}_3(\boldsymbol{p}) = (\boldsymbol{h}_1(\boldsymbol{p})^t, \boldsymbol{h}_2(\boldsymbol{p})^t)^t.$$

Note that $\boldsymbol{h}_s(\boldsymbol{p})$, $s = 1, 2, 3$, are the vectors of order $d_s \times 1$, and d_s, $s = 1, 2, 3$, are the numbers of df for testing goodness-of-fit of the QP$[f]_2^3$, MP$_2^3$ and PS^3 models, respectively. Therefore, we can prove Theorem 5.2 when r is odd, in a similar manner to the case of $T = 2$. The other cases can be proved in a similar way. The proof is complete. □

Finally, for an r^T table, we obtain the following theorem.

Theorem 5.3. *For an r^T table and a fixed h ($h = 1, \ldots, T-1$), $G^2(PS^T)$ is asymptotically equivalent to the sum of $G^2(QP[f]_h^T)$ and $G^2(MP_h^T)$ under the PS^T model.*

The proof of Theorem 5.3 is omitted because it can be obtained in a similar way as the proof of Theorem 5.2. We note that the theorem with $f(u) = u \log u$ is given by Tahata and Tomizawa (2008).

§6. Examples

6.1. Example 1

Consider the data in Table 2, taken from Tomizawa (1985), that is constructed from the data of the unaided distance vision of 4746 students aged 18 to about 25 including about 10% women in Faculty of Science and Technology, Science

University of Tokyo in Japan examined in April 1982. The row and column variables are the right and left eye grades, respectively, with the categories ordered from Best grade (1) to Worst grade (4).

We set $f(u) = u \log u$ and $f(u) = (1-u)^2$ applying the QP[f] model to the data. Table 3 gives the values of likelihood ratio chi-square statistics G^2 for testing the goodness-of-fit of models applied to the data in Table 2. The PS and MP models fit the data poorly, whereas the QP and Pearsonian-QP models fit the data well. For example, under the Pearsonian-QP model, it is inferred that there is the structure of point-symmetry in terms of $\theta^P_{(i<j;k<l)}$. We note that $G^2(PS)$ is close to the sum of $G^2(QP)$ (or $G^2(Pearsonian\text{-}PS)$) and $G^2(MP)$.

6.2. Example 2

Consider the data in Table 4, taken from Tahata, Tokuno and Tomizawa (2010). These data are obtained from the Meteorological Agency in Japan. These are obtained from the daily temperatures at Sapporo City, Japan, in three years, 2001, 2002 and 2003, using three levels, (1) below normals, (2) normals and (3) above normals.

We set $f(u) = u \log u$ and $f(u) = (1-u)^2$ applying the QP$[f]^3_1$ and QP$[f]^3_2$ models to the data. Table 5 gives the values of likelihood ratio chi-square statistics G^2 for testing the goodness-of-fit of models applied to the data in Table 4. Table 5 shows that the PS^3, MP$^3_1$ and MP$^3_2$ models fit the data poorly, but the QP$^3_1$, QP$^3_2$, Pearsonian-QP$^3_1$ and Pearsonian-QP$^3_2$ models fit the data very well. Consider the hypothesis that the QP$^3_1$ (or Pearsonian-QP$^3_1$) model holds assuming that the QP$^3_2$ (or Pearsonian-QP$^3_2$) model holds. Since $G^2(QP^3_1 \mid QP^3_2) = 8.44$ (or $G^2(Pearsonian\text{-}QP^3_1 \mid Pearsonian\text{-}QP^3_2) = 8.44$) with 6 df, we accept the hypothesis at the 0.05 level. Therefore, the QP$^3_1$ (or Pearsonian-QP$^3_1$) model may be preferable to the QP$^3_2$ (or Pearsonian-QP$^3_2$) model. Hence, for example, it is inferred that there is the structure of point-symmetry in terms of equation (3.1) with $f(u) = (1-u)^2$. We note that $G^2(PS^3)$ is close to (i) the sum of $G^2(QP^3_1)$ (or $G^2(Pearsonian\text{-}QP^3_1)$) and $G^2(MP^3_1)$, and (ii) the sum of $G^2(QP^3_2)$ (or $G^2(Pearsonian\text{-}QP^3_2)$) and $G^2(MP^3_2)$.

§7. Concluding remarks

In this paper, we have proposed a model, which is the closest model to the PS model in terms of the f-divergence under the condition that row marginals $\{p_{i\cdot}\}$, column marginals $\{p_{\cdot j}\}$ and sums $\{p_{ij} + p_{i^*j^*}\}$ are given. The proposed model has its characterization in terms of point-symmetry of $\theta^{[f]}_{(i<j;k<l)}$. We note that the QP[f] model indicates the point-symmetry of $\theta_{(i<j;k<l)}$ when we set $f(u) = u \log u$, whereas the QP[f] model indicates the point-symmetry of $\theta^{P}_{(i<j;k<l)}$ when we set $f(u) = (1-u)^2$. We note that $\theta^{[f]}_{(i<j;k<l)}$ reduces to $\theta_{(i<j;k<l)}$ when we set $f(u) = u \log u$, whereas $\theta^{[f]}_{(i<j;k<l)}$ reduces to $\theta^{P}_{(i<j;k<l)}$ when we set $f(u) = (1-u)^2$.

Moreover, we have given the decomposition of the PS model into the QP[f] and MP models. This theorem includes the decomposition of the PS model given by Tomizawa (1985) as a special case. Here we consider the artificial data in Table 6. We set $f(u) = u \log u$ and $f(u) = (1-u)^2$ applying the QP[f] model to the data. Table 7 gives the values of likelihood ratio chi-square statistics G^2 for testing the goodness-of-fit of models applied to the data in Table 6. These show that the Pearsonian-QP model fits the data well although the other models including the QP model fit the data poorly. Therefore, Theorem 2.2 indicates that the poor fit of the PS model is due to the lack of structure of the MP model rather than the Pearsonian-QP model. Since both the QP and MP models fit the data poorly, existing decomposition is not able to reveal the origin of the poor fit of the PS model for these data. Hence, it may be possible to explain the reason of poor fit of the PS model for more details when it occurs for a real dataset by using our proposed decomposition. Also, we have shown the property of test statistics for the decomposition of the PS model into the QP[f] and MP models. Furthermore, we extended the QP[f] model and decomposition theorem for the PS model into multi-way tables.

Acknowledgments

The authors are greatly thankful to the referee and editor for their helpful and valuable suggestions to further improve this article.

References

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

[2] Csiszár, I. and Shields, P. C. (2004). *Information Theory and Statistics: A Tutorial.* Now Publishers, Hanover, Massachusetts.

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

[4] Gilula, Z., Krieger, A. M. and Ritov, Y. (1988). Ordinal association in contingency tables: some interpretive aspects. *Journal of the American Statistical Association*, **83**, 540-545.

[5] Kateri, M. (2018). ϕ-divergence in contingency table analysis. *Entropy*, **20**, 1-12.

[6] Kateri, M. and Agresti. A. (2007). A class of ordinal quasi-symmetry models for square contingency tables. *Statistics and Probability Letters*, **77**, 598-603.

[7] Kateri, M. and Papaioannou, T. (1997). Asymmetry models for contingency tables. *Journal of the American Statistical Association*, **92**, 1124-1131.

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

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

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

[11] Saigusa, Y., Tahata, K. and Tomizawa, S. (2015). Orthogonal decomposition of symmetry model using the ordinal quasi-symmetry model based on f-divergence for square contingency tables. *Statistics and Probability Letters*, **101**, 33-37.

[12] Tahata, K., Tokuno, H. and Tomizawa, S. (2010). On test of marginal point-symmetry in multi-way tables. *Advances and Applications in Statistical Sciences*, **3**, 1-12.

[13] Tahata, K. and Tomizawa, S. (2008). Orthogonal decomposition of point-symmetry for multi-way tables. *Advances in Statistical Analysis*, **92**, 255-269.

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

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

[16] Tomizawa, S. and Tahata, K. (2007). The analysis of symmetry and asymmetry: orthogonality of decomposition of symmetry into quasi-symmetry and marginal symmetry for multi-way tables. *Journal de la Société Française de Statistique*, **148**, 3-36.

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

Table 1. Numbers of degrees of freedom (df) for models applied to the r^T table.

Models	df	
	When r is even	When r is odd
PS^T	$\dfrac{r^T}{2}$	$\dfrac{r^T-1}{2}$
$\mathrm{QP}[f]_h^T$	$\dfrac{r^T}{2} - \sum_{k=1}^{h} \binom{T}{k} C_k$	$\dfrac{r^T-1}{2} - \sum_{k=1}^{h} \binom{T}{k} \dfrac{(r-1)^k}{2}$
MP_h^T	$\sum_{k=1}^{h} \binom{T}{k} C_k$	$\sum_{k=1}^{h} \binom{T}{k} \dfrac{(r-1)^k}{2}$

Note:

$$\binom{T}{k} = \frac{T!}{k!(T-k)!}, \quad C_k = \begin{cases} \dfrac{(r-1)^k+1}{2} & (k : \text{odd}), \\ \dfrac{(r-1)^k-1}{2} & (k : \text{even}). \end{cases}$$

Table 2. Unaided distance vision of 4746 students aged 18 to about 25 including about 10% women in Faculty of Science and Technology, Science University of Tokyo in Japan examined in April 1982; adapted from Tomizawa (1985).

Right eye grade	Left eye grade				Total
	Best(1)	Second(2)	Third(3)	Worst(4)	
Best(1)	1291	130	40	22	1483
	(1306.46)[a]	(122.65)	(36.90)	(17.00)	
	(1306.21)[b]	(124.27)	(36.39)	(17.11)	
Second(2)	149	221	114	23	507
	(134.33)	(227.56)	(116.31)	(28.80)	
	(136.03)	(224.58)	(115.87)	(29.10)	
Third(3)	64	124	660	185	1033
	(58.20)	(121.69)	(653.44)	(199.67)	
	(57.90)	(122.13)	(656.42)	(197.97)	
Worst(4)	20	25	249	1429	1723
	(25.00)	(28.10)	(256.35)	(1413.54)	
	(24.89)	(28.61)	(254.73)	(1413.79)	
Total	1524	500	1063	1659	4746

Note: [a]MLEs under the QP model.
[b]MLEs under the Pearsonian-QP model.

Table 3. Values of the likelihood ratio chi-squared statistic (G^2) for models applied to Table 2.

Models	df	G^2
PS	8	301.86*
QP	4	8.84
Pearsonian-QP	4	8.07
MP	4	293.79*

*significant at the 0.05 level.

Table 4. The daily atmospheric temperatures at Sapporo City, Japan, in 2001, 2002 and 2003, using three levels, (1) below normals, (2) normals and (3) above normals; from Tahata *et al.* (2010).

2003	2001	2002		
		(1)	(2)	(3)
(1)	(1)	11	10	12
		(11.24)[a]	(13.26)	(10.25)
		(9.18)[b]	(11.74)	(12.08)
		(11.29)[c]	(13.29)	(10.27)
		(9.19)[d]	(11.79)	(12.17)
(1)	(2)	12	15	17
		(8.99)	(13.53)	(19.62)
		(12.39)	(13.07)	(18.54)
		(8.97)	(13.54)	(19.60)
		(12.38)	(13.07)	(18.56)
(1)	(3)	5	12	12
		(7.70)	(13.13)	(8.28)
		(6.43)	(12.19)	(10.38)
		(7.58)	(13.10)	(8.25)
		(6.30)	(12.14)	(10.39)
(2)	(1)	15	19	15
		(15.28)	(16.85)	(15.23)
		(15.20)	(17.45)	(16.35)
		(15.34)	(16.87)	(15.30)
		(15.24)	(17.36)	(16.24)
(2)	(2)	13	23	12
		(10.17)	(23.00)	(14.83)
		(14.15)	(23.00)	(10.85)
		(10.17)	(23.00)	(14.83)
		(14.16)	(23.00)	(10.84)
(2)	(3)	9	12	19
		(8.77)	(14.15)	(18.72)
		(7.65)	(13.55)	(18.80)
		(8.70)	(14.13)	(18.66)
		(7.76)	(13.64)	(18.76)

Note: [a]MLEs under the QP_1^3 model.
[b]MLEs under the QP_2^3 model.
[c]MLEs under the Pearsonian-QP_1^3 model.
[d]MLEs under the Pearsonian-QP_2^3 model.

Table 4. (continued)

2003	2001	2002		
		(1)	(2)	(3)
(3)	(1)	4	19	18
		(7.72)[a]	(17.87)	(15.30)
		(5.62)[b]	(18.81)	(16.57)
		(7.75)[c]	(17.90)	(15.42)
		(5.61)[d]	(18.86)	(16.70)
(3)	(2)	18	14	12
		(15.38)	(15.47)	(15.01)
		(16.46)	(15.93)	(11.61)
		(15.40)	(15.46)	(15.03)
		(16.44)	(15.93)	(11.62)
(3)	(3)	5	16	16
		(6.75)	(12.74)	(15.76)
		(4.92)	(14.26)	(17.82)
		(6.73)	(12.71)	(15.71)
		(4.83)	(14.21)	(17.81)

Note: [a]MLEs under the QP_1^3 model.
[b]MLEs under the QP_2^3 model.
[c]MLEs under the Pearsonian-QP_1^3 model.
[d]MLEs under the Pearsonian-QP_2^3 model.

Table 5. Values of the likelihood ratio chi-squared statistic (G^2) for models applied to Table 4.

Models	df	G^2
PS^3	13	22.56*
QP_1^3	10	12.37
QP_2^3	4	3.93
Pearsonian-QP_1^3	10	12.30
Pearsonian-QP_2^3	4	3.86
MP_1^3	3	10.27*
MP_2^3	9	18.70*

*significant at the 0.05 level.

Table 6. Artificial data.

X_1	X_2 (1)	(2)	(3)	Total
(1)	40	110	350	500
	(39.78)[a]	(117.65)	(348.30)	
(2)	20	60	120	200
	(17.00)	(60.00)	(123.00)	
(3)	21	100	179	300
	(22.70)	(92.35)	(179.22)	
Total	81	270	649	1000

Note: [a]MLEs under the Pearsonian-QP model.

Table 7. Values of the likelihood ratio chi-squared statistic (G^2) for models applied to Table 6.

Models	df	G^2
PS	4	528.02*
QP	2	9.48*
Pearsonian-QP	2	1.84
MP	2	526.18*

*significant at the 0.05 level.

Takuya Yoshimoto
Department of Information Sciences, Faculty of Science and Technology,
Tokyo University of Science
Noda City, Chiba, 278-8510, Japan
Clinical Information & Intelligence Dept., Chugai Pharmaceutical Co., Ltd.
Chuo-ku, Tokyo, 103-8324, Japan
E-mail: yoshimoto.takuya61@chugai-pharm.co.jp

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

Yusuke Saigusa
Department of Biostatistics, Yokohama City University School of Medicine
Yokohama City, Kanagawa, 236-0004, Japan
E-mail: saigusay@yokohama-cu.ac.jp

Sadao Tomizawa
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