HIROSHIMA MATH. J.
31 (2001), 1–34

Tests for redundancy of some variables in correspondence analysis

Teruyuki NAKAYAMA

(Received March 24, 2000)
(Revised May 17, 2000)

ABSTRACT. This paper deals with the problems of formulating and testing the hypotheses of redundancy of some variables in correspondence analysis for a two-way contingency table. The testing problems are examined under the assumption that the contingency table is multinomial or independent multinomial. The asymptotic properties of the MLE's and the test statistics under the null hypotheses are also examined.

1. Introduction

This paper is concerned with correspondence analysis (abbreviated as CA), which has been extended as a method of scaling the categories of a two-way contingency table with r rows, c columns and n_{ij} observations in the (i, j)-th cell. Suppose that the data are expressed in a table as in Figure 1. Then, we call the categories $A_1, A_2, \dots, A_r$ and the categories $B_1, B_2, \dots, B_c$ as the row variables and the column variables, respectively. It is regarded that the i-th row variable has c-dimensional profile $(n_{i1}/n_{i\cdot}, \dots, n_{ic}/n_{i\cdot})$, and the j-th column variable has r-dimensional profile $(n_{1j}/n_{\cdot j}, \dots, n_{rj}/n_{\cdot j})$, where $n_{i\cdot} = \sum_{j=1}^{c} n_{ij}$ and $n_{\cdot j} = \sum_{i=1}^{r} n_{ij}$. The object of CA is to scale the row variables and the column variables, or more precisely, to give a simultaneous low dimensional plot of the row variables and the column variables in the table. In this plot, it is intended that variables that are close to each other in the plot will be the ones that are closely related to each other according to the profiles. Hill [8] discussed a history of its development. Recently, the asymptotic distributions of some basic statistics related to CA were given by Eaton and Tyler [2]. Rao [17, 18] developed a general theory of canonical variate methods including the usual CA. Relating to Rao's work, Nakayama *et al.* [12] examined stabilities of the configurations obtained from the usual CA and the canonical variate method.

On the other hand, Rao [13, 14, 15] formulated the notation of redundancy of a given set of variables and studied the testing problem in discriminant analysis. Since then, the idea has been extended to various mul-

2000 *Mathematics Subject Classification*. 62H25, 62H17
Key words and phrases. asymptotic distribution, contingency table, multinomial distribution.

	B_1	B_2	$\dots$	B_c	Total
A_1	n_{11}	n_{12}	$\cdots$	n_{1c}	$n_{1\cdot}$
A_2	n_{21}	n_{22}	$\cdots$	n_{2c}	$n_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
A_r	n_{r1}	n_{r2}	$\cdots$	n_{rc}	$n_{r\cdot}$
Total	$n_{\cdot 1}$	$n_{\cdot 2}$	$\cdots$	$n_{\cdot c}$	n

Fig. 1. Data for a two-way contingency table.

tivariate situations. Further, Akaike's information criterion has been derived for such hypotheses or models of redundancy (see e.g., Fujikoshi [3, 4]).

The main object of the present paper is to examine the problems of formulating and testing the hypotheses for redundancy of some variable in CA. The paper is organized in the following way: In §2, we describe the basic method of CA. In §3, by considering an additional column variable, two types of redundancies are formulated along an aim of CA. One is the redundancy of an additional column variable for scaling of the row variables. The other is the redundancy of an additional column variable for scaling of the column variables. These formulations are applied for three models of the contingency table. In §4, the testing problems of redundancy of a column variable for scaling of the row variables are studied for each model. These problems are corresponding to the ones for redundancy of a variable in discriminant analysis. In §5, the testing problems of redundancy of a column variable for scaling of the column variables are studied for each model. In §6, we give examples and numerical experiments that illustrate our results for each case. In §7, the proofs of theorems are given.

2. Correspondence analysis

Let $\boldsymbol{N}$ be an $r \times c$ data matrix related to a two-way table as in Figure 1, $\boldsymbol{F} = n^{-1}\boldsymbol{N}$, $\boldsymbol{D}_r$ an $r \times r$ diagonal matrix with the i-th diagonal element $n_{i\cdot}/n$, and $\boldsymbol{D}_c$ a $c \times c$ diagonal matrix with the j-th diagonal element $n_{\cdot j}/n$. In CA of the two-way contingency table (see e.g., Lebart *et al.* [11], Greenacre [6], etc.), the squared distance between the j-th column variable and the j'-th column variable is measured by the chi-square distance defined as

$$d_c^2(j, j') = \sum_{i=1}^{r} \frac{n}{n_{i\cdot}} \left(\frac{n_{ij}}{n_{\cdot j}} - \frac{n_{ij'}}{n_{\cdot j'}} \right)^2. \tag{1}$$

Similarly, the chi-square distance between the i-th row variable and the i'-th row variable is measured by

$$d_r^2(i,i') = \sum_{j=1}^{c} \frac{n}{n_{\cdot j}} \left(\frac{n_{ij}}{n_{i\cdot}} - \frac{n_{i'j}}{n_{i'\cdot}} \right)^2. \tag{2}$$

Note that the distance $d_c(j,j')$ can be regarded as the Euclidean distance when the j-th column variable has the coordinate in an r-dimensional Euclidean space (E^r) given by the j-th column vector of $\boldsymbol{D}_r^{-1/2}\boldsymbol{F}\boldsymbol{D}_c^{-1}$. Similarly the distance $d_r(i,i')$ can be regarded as the Euclidean distance when the i-th row variable has the coordinate in a c-dimensional Euclidean space given by the i-th column vector of $\boldsymbol{D}_c^{-1/2}\boldsymbol{F}'\boldsymbol{D}_r^{-1}$. Now we consider to present the column variables in E^r as the point in E^k ($k < r$), in such a way that the relative positions of the column variables in E^r are preserved as the extent possible in E^k.

Let a $c \times k$ matrix $\boldsymbol{Y}$ be the configuration matrix of the column variables in the reduced space E^k, *i.e.*, the j-th row of $\boldsymbol{Y}$ denotes the coordinates for the configuration of the j-th column variable in E^k. Then, our problem is to find $\boldsymbol{Y}$ to minimizing

$$\|\boldsymbol{D}_c^{-1}\boldsymbol{F}'\boldsymbol{D}_r^{-1}\boldsymbol{F}\boldsymbol{D}_c^{-1} - \boldsymbol{Y}\boldsymbol{Y}'\| \tag{3}$$

with respect to $\boldsymbol{Y}$. The solution is given as follows (see, Rao [17, 18]).

Consider the singular value decomposition (*s.v.d.*)

$$\boldsymbol{D}_r^{-1/2}\boldsymbol{F}\boldsymbol{D}_c^{-1/2} = \sqrt{\ell_0}\boldsymbol{v}_0\boldsymbol{w}_0' + \sqrt{\ell_1}\boldsymbol{v}_1\boldsymbol{w}_1' + \cdots + \sqrt{\ell_K}\boldsymbol{v}_K\boldsymbol{w}_K',$$

where $1 = \sqrt{\ell_0} \geq \sqrt{\ell_1} \geq \cdots \geq \sqrt{\ell_K} > 0$, $\boldsymbol{v}_0 = \boldsymbol{D}_r^{1/2}\mathbf{1}_r$, $\boldsymbol{w}_0 = \boldsymbol{D}_c^{1/2}\mathbf{1}_c$, $K+1 = \operatorname{rank}(\boldsymbol{D}_r^{-1/2}\boldsymbol{F}\boldsymbol{D}_c^{-1/2}) \leq \min\{r,c\}$ *and* $\mathbf{1}_r$ *is the r-dimensional vector of unities. Note that the vectors satisfy the restrictions* $\boldsymbol{v}_\alpha'\boldsymbol{v}_\beta = \boldsymbol{w}_\alpha'\boldsymbol{w}_\beta = \delta_{\alpha\beta}$, *where Kronecker's* $\delta_{\alpha\beta}$ *is* 1 *if* $\alpha = \beta$, 0 *if* $\alpha \neq \beta$. *Then the choice*

$$\boldsymbol{Y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k) = (\sqrt{\ell_1}\boldsymbol{D}_c^{-1/2}\boldsymbol{w}_1, \ldots, \sqrt{\ell_k}\boldsymbol{D}_c^{-1/2}\boldsymbol{w}_k)$$

minimizes (3).

From the above s.v.d., it is seen that a "best" approximation matrix with rank k ($k \leq K$) to $\boldsymbol{D}_r^{-1/2}\boldsymbol{F}\boldsymbol{D}_c^{-1}$ is expressed as

$$\boldsymbol{v}_0\boldsymbol{y}_0' + \boldsymbol{v}_1\boldsymbol{y}_1' + \cdots + \boldsymbol{v}_k\boldsymbol{y}_k' = \boldsymbol{v}_0\boldsymbol{y}_0' + [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k]\boldsymbol{Y}'.$$

This result gives a justification for presenting the j-th column variable as a point in the reduced space E^k with the coordinate given by the j-th column vector of $\boldsymbol{Y}$.

Similarly, the configuration matrix $r \times k$ matrix $\boldsymbol{X}$ of the row variables in the reduced space E^k is obtained by minimizing

$$\|\boldsymbol{D}_r^{-1}\boldsymbol{F}\boldsymbol{D}_c^{-1}\boldsymbol{F}'\boldsymbol{D}_r^{-1} - \boldsymbol{X}\boldsymbol{X}'\|$$

with respect to $\boldsymbol{X}$. In terms of the s.v.d of $\boldsymbol{D}_c^{-1/2}\boldsymbol{F}'\boldsymbol{D}_r^{-1/2}$, we can write $\boldsymbol{X}$ as

$$\boldsymbol{X} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_k) = (\sqrt{\ell_1}\boldsymbol{D}_r^{-1/2}\boldsymbol{v}_1, \dots, \sqrt{\ell_k}\boldsymbol{D}_r^{-1/2}\boldsymbol{v}_k).$$

Such a configuration is also justified as a "best" approximation matrix with rank k $(k \le K)$ to $\boldsymbol{D}_c^{-1/2}\boldsymbol{F}'\boldsymbol{D}_r^{-1}$ which is expressed as

$$\boldsymbol{w}_0\boldsymbol{x}_0' + \boldsymbol{w}_1\boldsymbol{x}_1' + \cdots + \boldsymbol{w}_k\boldsymbol{x}_k' = \boldsymbol{w}_0\boldsymbol{x}_0' + [\boldsymbol{w}_1, \dots, \boldsymbol{w}_k]\boldsymbol{X}'.$$

Note that, since $\sqrt{\ell_\alpha}\boldsymbol{v}_\alpha$ may be denoted as $\boldsymbol{D}_r^{-1/2}\boldsymbol{F}\boldsymbol{D}_c^{-1/2}\boldsymbol{w}_\alpha$, the configuration matrix $\boldsymbol{X}$ can be rewritten as

$$\boldsymbol{X} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_k) = (\boldsymbol{D}_r^{-1}\boldsymbol{F}\boldsymbol{D}_c^{-1/2}\boldsymbol{w}_1, \dots, \boldsymbol{D}_r^{-1}\boldsymbol{F}\boldsymbol{D}_c^{-1/2}\boldsymbol{w}_k).$$

As both the configuration matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ are expressed in terms of $\boldsymbol{w}_\alpha$, they may be also simultaneously represented as points on the spaces spanned by $\boldsymbol{w}_1, \dots, \boldsymbol{w}_k$.

In the following, we will use alternative expressions for $\boldsymbol{Y}$ and $\boldsymbol{X}$ in terms of the eigenvalue-eigenvector problem of $\boldsymbol{S}$ in (4).

Consider the eigenvalue-eigenvector problem of

$$\boldsymbol{S} = \boldsymbol{D}_c^{-1}\boldsymbol{F}'\boldsymbol{D}_r^{-1}\boldsymbol{F} \tag{4}$$

with eigenvalues $\ell_0 \ge \ell_1 \ge \cdots \ge \ell_{c-1} \ge 0$ *and eigenvectors* $\boldsymbol{u}_0, \boldsymbol{u}_1, \dots, \boldsymbol{u}_{c-1}$, *i.e.,* $\boldsymbol{S}\boldsymbol{u}_\alpha = \ell_\alpha\boldsymbol{u}_\alpha$ $(\alpha = 0, \dots, c-1)$ *with the restriction* $\boldsymbol{u}_\alpha'\boldsymbol{D}_c\boldsymbol{u}_\beta = \delta_{\alpha\beta}$, *where* $\ell_0 = 1$ *and* $\boldsymbol{u}_0 = \boldsymbol{1}_c$. *Then the configuration matrix* $\boldsymbol{Y}$ *can be expressed as*

$$\boldsymbol{Y} = (\boldsymbol{y}_1, \dots, \boldsymbol{y}_k) = (\sqrt{\ell_1}\boldsymbol{u}_1, \dots, \sqrt{\ell_k}\boldsymbol{u}_k). \tag{5}$$

Similarly, $\boldsymbol{X}$ *is expressed as*

$$\boldsymbol{X} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_k) = (\boldsymbol{D}_r^{-1}\boldsymbol{F}\boldsymbol{u}_1, \dots, \boldsymbol{D}_r^{-1}\boldsymbol{F}\boldsymbol{u}_k). \tag{6}$$

3. Formulation of redundancy

It is important to examine an influence of a column variable for scaling of the row or the column variables. In this section, we investigate the problem of formulating whether a column variable is redundant for scaling of the row or the column variables. Similarly, we can treat the problem of formulating whether a row variable is redundant for scaling of the column or the row variables. For simplicity, we consider an additional column variable which has an r-dimensional profile $\boldsymbol{m} = (m_1/m, \dots, m_r/m)'$, where $m = \sum_{i=1}^{r} m_i$. Without loss of generality, we assume that an additional column variable is the $(c+1)$-th column variable. We consider the problem of formulating whether the $(c+1)$-th column variable is redundant or not. First we prepare some notations. Let

$$\tilde{\boldsymbol{N}} = [\boldsymbol{N}|\boldsymbol{m}] : r \times (c+1), \qquad \tilde{\boldsymbol{F}} = \frac{1}{n+m}\tilde{\boldsymbol{N}} = \left[\frac{n}{n+m}\boldsymbol{F} \,\middle|\, \frac{m}{n+m}\boldsymbol{f}_m\right],$$

$$\tilde{\boldsymbol{D}}_r = \frac{n}{n+m}\boldsymbol{D}_r + \frac{m}{n+m}\boldsymbol{D}_M, \qquad \tilde{\boldsymbol{D}}_c = \left[\begin{array}{c|c} \frac{n}{n+m}\boldsymbol{D}_c & \boldsymbol{0} \\ \hline \boldsymbol{0}' & \frac{m}{n+m} \end{array}\right],$$

where $\boldsymbol{f}_m = \frac{1}{m}\boldsymbol{m}$, $\boldsymbol{D}_M = \mathrm{diag}(m_1/m, \ldots, m_r/m)$. Then the $(c+1) \times (c+1)$ matrix corresponding to (4) is denoted as $\tilde{\boldsymbol{S}} = \tilde{\boldsymbol{D}}_c^{-1}\tilde{\boldsymbol{F}}'\tilde{\boldsymbol{D}}_r^{-1}\tilde{\boldsymbol{F}}$ with eigenvalues $1 = \tilde{\ell}_0 \geq \tilde{\ell}_1 \geq \cdots \geq \tilde{\ell}_c \geq 0$, eigenvectors $\tilde{\boldsymbol{u}}_0, \tilde{\boldsymbol{u}}_1, \ldots, \tilde{\boldsymbol{u}}_c$ and $\tilde{\boldsymbol{u}}_0 = \boldsymbol{1}_{c+1}$. Let $\tilde{k} + 1 = \mathrm{rank}(\tilde{\boldsymbol{S}}) \leq \min\{r, c+1\}$. The vectors corresponding to $\boldsymbol{y}_\alpha$ in (5) and $\boldsymbol{x}_\alpha$ in (6) are denoted as

$$\tilde{\boldsymbol{y}}_\alpha = \sqrt{\tilde{\ell}_\alpha}\tilde{\boldsymbol{u}}_\alpha, \qquad \tilde{\boldsymbol{x}}_\alpha = \tilde{\boldsymbol{D}}_r^{-1}\tilde{\boldsymbol{F}}\tilde{\boldsymbol{u}}_\alpha, \qquad \alpha = 1, \ldots, \tilde{k},$$

respectively. That is, the α-th coordinates for the j-th column variable and the i-th row variable are denoted as

$$\tilde{y}_{j\alpha} = \sqrt{\tilde{\ell}_\alpha}\tilde{u}_{j\alpha}, \qquad \tilde{x}_{i\alpha} = \frac{1}{n_{i\cdot} + m_i}\left\{\sum_{j=1}^{c} n_{ij}\tilde{u}_{j\alpha} + m_i\tilde{u}_{(c+1)\alpha}\right\},$$

respectively.

Then, the chi-square distance between the i-th row variable and the i'-th row variable corresponding to (2) is denoted as

$$\tilde{d}_r^2(i, i') = \sum_{j=1}^{c}\frac{n+m}{n_{\cdot j}}\left(\frac{n_{ij}}{n_{i\cdot} + m_i} - \frac{n_{i'j}}{n_{i'\cdot} + m_{i'}}\right)^2 + \frac{n+m}{m}\left(\frac{m_i}{n_{i\cdot} + m_i} - \frac{m_{i'}}{n_{i'\cdot} + m_{i'}}\right)^2.$$

It is natural to say that if $\tilde{d}_r^2(i, i') = \kappa d_r^2(i, i')$ $(\kappa > 0)$ for any i and i', the $(c+1)$-th column variable is redundant for scaling of the row variables. In fact, the condition $\tilde{d}_r^2(i, i') = \kappa d_r^2(i, i')$ for any i and i' is equivalent to the condition $\tilde{x}_{i\alpha} = \kappa x_{i\alpha}$ for any i and $\alpha = 1, \ldots, k$. Further, each of them implies $\tilde{u}_{j\alpha} = u_{j\alpha}$ and $k = \tilde{k}$. This means that the configuration of the row variables based on the $r \times c$ contingency table is essentially the same as the one based on the $r \times (c+1)$ contingency table. These results are proved, in terms of population parameters, in §7.

On the other hand, the chi-square distance between a t-th column variable and the $(c+1)$-th column variable corresponding to (1) is denoted as

$$\tilde{d}_c^2(t, c+1) = \sum_{i=1}^{r}\frac{n+m}{n_{i\cdot} + m_i}\left(\frac{n_{it}}{n_{\cdot t}} - \frac{m_i}{m}\right)^2.$$

If $\tilde{d}_c^2(t, c+1) = 0$, *i.e.*, the profile of the $(c+1)$-th column variable is the same as the one of a t-th column variable, we have $\tilde{d}_c^2(j, t) = \tilde{d}_c^2(j, c+1)$ for any j, and hence we can see that the $(c+1)$-th column variable does not give any influence in scaling of the column variables. More precisely, we will see that if there is some $t \in \{1, \dots, c\}$ such that $\tilde{d}_c^2(t, c+1) = 0$, the $(c+1)$-th column variable is redundant for scaling of the column variables.

Based on the above consideration, we introduce a notion of redundancy of the $(c+1)$-th column variable in terms of population parameters. A general model for two-way contingency table is described as the conditional Poisson model; see Haberman [7]. In this paper, for an $r \times (c+1)$ random matrix $\tilde{\boldsymbol{N}} = [\boldsymbol{N}|\boldsymbol{m}]$, we treat the following three conditional Poisson models which are also described as the variants of multinomial distributions.

Model (1):
Assume that $(n+m)$ is given. The entries of $\tilde{\boldsymbol{N}}$ are jointly multinomial with probabilities $\tilde{\boldsymbol{P}}$ and trial size parameter $(n+m) > 0$, where

$$\tilde{\boldsymbol{P}} = [\boldsymbol{P}|\boldsymbol{q}] = \left[\begin{array}{ccc|c} p_{11} & \cdots & p_{1c} & q_1 \\ \vdots & \ddots & \vdots & \vdots \\ p_{r1} & \cdots & p_{rc} & q_r \end{array}\right].$$

Model (2):
Assume that $n_{i\cdot} + m_i$ $(i = 1, \dots, r)$ are given. The rows of $\tilde{\boldsymbol{N}}$, *i.e.* $\boldsymbol{n}_i = (n_{i1}, \dots, n_{ic}, m_i)'$ have r independent multinomials with respective probability vectors $\boldsymbol{p}_i = (p_{i1}, \dots, p_{ic}, q_i)'$ and trial size parameters $(n_{i\cdot} + m_i)$.

Model (3):
Assume that $n_{\cdot j}$ $(j = 1, \dots, c)$ and m are given. The columns of $\tilde{\boldsymbol{N}}$, *i.e.* $\boldsymbol{n}_j = (n_{1j}, \dots, n_{rj})'$ and $\boldsymbol{m}$ have $(c+1)$ independent multinomials with respective probability vectors $\boldsymbol{p}_j = (p_{1j}, \dots, p_{rj})'$, $\boldsymbol{q}$ and trial size parameters $n_{\cdot j}$, m.

For these models, the eigenvalue-eigenvector problem of CA in terms of population parameters can be written as follows. Let $\boldsymbol{\Theta}$ be the matrix obtained from $\boldsymbol{S}$ by substituting p_{ij} for n_{ij}, $\tilde{\boldsymbol{\Theta}}$ the matrix obtained from $\tilde{\boldsymbol{S}}$ by substituting p_{ij} for n_{ij} and q_i for m_i, respectively. Also, let $\tau_r^2(i, i')$, $\tilde{\tau}_r^2(i, i')$ and $\tilde{\tau}_c^2(j, c+1)$ be the chi-square distance obtained from $d_r^2(i, i')$, $\tilde{d}_r^2(i, i')$ and $\tilde{d}_c^2(j, c+1)$ by substituting p_{ij} for n_{ij} and q_i for m_i, respectively. Let $d - 1 = \min(r, c)$ and $\tilde{d} - 1 = \min(r, c+1)$. Assume that $k = \operatorname{rank}(\boldsymbol{\Theta}) - 1$ and $\tilde{k} = \operatorname{rank}(\tilde{\boldsymbol{\Theta}}) - 1$. Let $\boldsymbol{\Theta}$ and $\tilde{\boldsymbol{\Theta}}$ have the eigenvalues and the eigenvectors expressed as

$$\boldsymbol{\Theta}\,\boldsymbol{\mu}_\alpha = \lambda_\alpha \boldsymbol{\mu}_\alpha, \qquad \boldsymbol{\mu}_\alpha' \boldsymbol{\Delta}_r \boldsymbol{\mu}_\beta = \delta_{\alpha\beta}, \qquad \tilde{\boldsymbol{\Theta}}\tilde{\boldsymbol{\mu}}_\alpha = \tilde{\lambda}_\alpha \tilde{\boldsymbol{\mu}}_\alpha, \qquad \tilde{\boldsymbol{\mu}}_\alpha' \tilde{\boldsymbol{\Delta}}_r \tilde{\boldsymbol{\mu}}_\beta = \delta_{\alpha\beta} \tag{7}$$

with $1 = \lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_k > \lambda_{k+1} = \cdots = \lambda_d = 0$ and $1 = \tilde{\lambda}_0 \geq \tilde{\lambda}_1 \geq \cdots \geq \tilde{\lambda}_{\tilde{k}} > \tilde{\lambda}_{\tilde{k}+1} = \cdots = \tilde{\lambda}_{\tilde{d}} = 0$. Here, $\boldsymbol{\Delta}_r = \mathrm{diag}(p_{i\cdot}/p_{\cdot\cdot})$, $\tilde{\boldsymbol{\Delta}}_r = \mathrm{diag}\{(p_{i\cdot}+q_i)/(p_{\cdot\cdot}+q_\cdot)\}$, $p_{i\cdot} = \sum_{j=1}^{c} p_{ij}$, $p_{\cdot\cdot} = \sum_{i=1}^{r}\sum_{j=1}^{c} p_{ij}$ and $q_\cdot = \sum_{i=1}^{r} q_i$. Note that we use the same notations k and $\tilde{k}$ as in case of sample observations. Then, for $\alpha = 1, \ldots, k$, the coordinates $\eta_{j\alpha}$ and $\xi_{i\alpha}$ corresponding to $y_{j\alpha}$ and $x_{i\alpha}$ are obtained by substituting p_{ij} for n_{ij} in $y_{j\alpha}$ and $x_{i\alpha}$, respectively, *i.e.*,

$$\eta_{j\alpha} = \sqrt{\lambda_\alpha}\mu_{j\alpha}, \qquad \xi_{i\alpha} = \frac{1}{p_{i\cdot}}\sum_{j=1}^{c} p_{ij}\mu_{j\alpha}.$$

For $\alpha = 1, \ldots, \tilde{k}$, the coordinates $\tilde{\eta}_{j\alpha}$ and $\tilde{\xi}_{i\alpha}$ corresponding to $\tilde{y}_{j\alpha}$ and $\tilde{x}_{i\alpha}$ are obtained by substituting p_{ij} for n_{ij} and q_i for m_i in $\tilde{y}_{j\alpha}$ and $\tilde{x}_{i\alpha}$, respectively, *i.e.*,

$$\tilde{\eta}_{j\alpha} = \sqrt{\tilde{\lambda}_\alpha}\tilde{\mu}_{j\alpha}, \tag{8}$$

$$\tilde{\xi}_{i\alpha} = \frac{1}{p_{i\cdot} + q_i}\left\{\sum_{j=1}^{c} p_{ij}\tilde{\mu}_{j\alpha} + q_i\tilde{\mu}_{(c+1)\alpha}\right\}. \tag{9}$$

As has been seen in case of sample observations, we have two types of redundancies of a column variable. One is the case that a column variable is redundant for scaling of the row variables, which is referred to as "Type (A)-redundancy". The other is the case that a column variable is redundant for scaling of the column variables, which is referred to as "Type (B)-redundancy". Now we give a formal definition of these redundancies of the $(c+1)$-th column variable in terms of population parameters, based on the distance, in the followings.

Definition A. *The $(c+1)$-th column variable is redundant for scaling of the row variables, if $\tilde{\tau}_r^2(i,i') = \phi\,\tau_r^2(i,i')$ for all i and i' $(i,i' = 1,\ldots,r)$ and some ϕ $(\phi > 0)$.*

Definition B. *The $(c+1)$-th column variable is redundant for scaling of the column variables, if there is some $t \in \{1,\ldots,c\}$ such that $\tilde{\tau}_c^2(t, c+1) = 0$.*

Note that without loss of generality, $0 < \phi < 1$. So, the condition in Definition A can be written as $\tilde{\tau}_r^2(i,i') = \tau_r^2(i,i')/(\kappa+1)$ $(\kappa > 0)$. The Definition A has the following equivalent statements. Proofs are all given in §7.

Theorem A. *For the statement that the $(c+1)$-th column variable is redundant for scaling of the row variables, the following statements are equivalent.*

(A1) $\tilde{\tau}_r^2(i,i') = \dfrac{1}{\kappa+1}\tau_r^2(i,i')$ *for* $i,i' = 1,\dots,r$ $(\kappa > 0)$,

(A2) $\tilde{\xi}_{i\alpha} = \dfrac{1}{\kappa+1}\xi_{i\alpha}$ *for* $i = 1,\dots,r$ *and* $\alpha = 1,\dots,k$,

(A3) $\tilde{\eta}_{(c+1)\alpha} = 0$ *for* $\alpha = 1,\dots,k$,

(A4) $\tilde{\mu}_{(c+1)\alpha} = 0$ *for* $\alpha = 1,\dots,k$,

(A5) $q_i = \kappa p_{i\cdot}$ $(\kappa > 0)$ *for* $i = 1,\dots,r$.

Further, each of the conditions (A1)–(A5) *implies* $k = \tilde{k}$ *and* $(\mu_{1\alpha},\dots,\mu_{c\alpha})' = (\tilde{\mu}_{1\alpha},\dots,\tilde{\mu}_{c\alpha})'$ *for* $\alpha = 1,\dots,k$.

Similarly, the Definition B has the following equivalent statements.

THEOREM B. *For the statement that the* $(c+1)$*-th column variable is redundant for scaling of the column variables, the following statements are equivalent.*

(B1) $\tilde{\tau}_c^2(t,c+1) = 0$ *for some* $t \in \{1,\dots,c\}$,

(B2) $\eta_{(c+1)\alpha} = \eta_{t\alpha}$ *for* $\alpha = 1,\dots,\tilde{k}$ *and some* $t \in \{1,\dots,c\}$,

(B3) $\mu_{(c+1)\alpha} = \mu_{t\alpha}$ *for* $\alpha = 1,\dots,\tilde{k}$ *and some* $t \in \{1,\dots,c\}$,

(B4) $\dfrac{q_i}{q_\cdot} = \dfrac{p_{it}}{p_{\cdot t}}$ *for* $i = 1,\dots,r$ *and some* $t \in \{1,\dots,c\}$.

In the case of Model (3)*, since* $p_{\cdot t} = q_\cdot = 1$*, the statement* (B4) *may be written as*

(B4*) $q_i = p_{it}$ *for* $i = 1,\dots,r$ *and some* $t \in \{1,\dots,c\}$.

These definitions and some statements in the theorems can be generalized for an $r \times c_1$ additional random matrix $\boldsymbol{M}$. For the entries of an $r \times (c+c_1)$ random matrix $[\boldsymbol{N}|\boldsymbol{M}]$, we consider three models corresponding to the Model (1), (2), (3). Let the probabilities corresponding to entries of $[\boldsymbol{N}|\boldsymbol{M}]$ be $[\boldsymbol{P}|\boldsymbol{Q}]$. Theorems A and B are generalized as Theorems A$'$ and B$'$, respectively.

THEOREM A$'$. *For the statement that the additional* c_1 *column variables are redundant for scaling of the row variables, we have the following equivalent statements.*

$$(\mathrm{A1}') \quad \tilde{\tau}_r^2(i,i') = \frac{1}{\sum_{j=1}^{c_1}\kappa_j + 1}\tau_r^2(i,i') \qquad \textit{for } i,i' = 1,\ldots,r,$$

$$(\mathrm{A2}') \quad q_{ij} = \kappa_j\, p_{i\cdot} \qquad \textit{for } i = 1,\ldots,r \textit{ and } j = 1,\ldots,c_1,$$

where κ_j's are some positive constants.

Theorem B$'$. *For the statement that the additional c_1 column variables are redundant for scaling of the column variables, we have the following equivalent statements.*

$$(\mathrm{B1}') \quad \tilde{\tau}_c^2(t_j, j) = 0 \qquad \textit{for some } t_j \in \{1,\ldots,c\}; \;\; j = 1,\ldots,c_1,$$

$$(\mathrm{B2}') \quad \frac{q_{ij}}{q_{\cdot j}} = \frac{p_{it_j}}{p_{\cdot t_j}} \qquad \textit{for } i = 1,\ldots,r; \;\; j = 1,\ldots,c_1.$$

In the case of Model (3), *the statement* (B2$'$) *may be written as*

$$(\mathrm{B2}'^{*}) \quad q_{ij} = p_{it_j} \qquad \textit{for } i = 1,\ldots,r; \;\; j = 1,\ldots,c_1.$$

4. Tests for Type (A)-redundancy of a column variable

Based on an $r \times (c+1)$ matrix $\tilde{\boldsymbol{N}}$ used in the previous section, we consider to test whether the $(c+1)$-th column variable is redundant for scaling of the row variables. Using the statement (A5) in Theorem A, we can express the testing problem as

$$\begin{aligned} &H_a : \frac{q_1}{p_{1\cdot}} = \cdots = \frac{q_r}{p_{r\cdot}}, \\ &K_a : \text{at least one strict inequality in } H_a. \end{aligned} \tag{10}$$

The testing problem is considered for three models (1), (2), (3).

4.1 Model (1)

Suppose that the entries of $\tilde{\boldsymbol{N}} = [\boldsymbol{N}|\boldsymbol{m}]$ are jointly multinomial with probabilities $\tilde{\boldsymbol{P}}$ and trial size parameter $(n+m)$, where

$$\tilde{\boldsymbol{P}} = [\boldsymbol{P}|\boldsymbol{q}] = \left[\begin{array}{ccc|c} p_{11} & \cdots & p_{1c} & q_1 \\ \vdots & \ddots & \vdots & \vdots \\ p_{r1} & \cdots & p_{rc} & q_r \end{array}\right],$$

$$\sum_{i=1}^{r}\sum_{j=1}^{c} p_{ij} + \sum_{i=1}^{r} q_i = 1. \tag{11}$$

The density function of $\tilde{\boldsymbol{N}}$ can be expressed as

$$f_1(\boldsymbol{N}, \boldsymbol{m}|\boldsymbol{P}, \boldsymbol{q}) = \theta_1 \prod_{i=1}^{r} \prod_{j=1}^{c} p_{ij}^{n_{ij}} \prod_{i=1}^{r} q_i^{m_i},$$

where $\theta_1 = (n+m)!/(\prod_{i=1}^{r} \prod_{j=1}^{c} n_{ij}! \prod_{i=1}^{r} m_i!)$. The maximum likelihood estimators (MLE) of p_{ij} and q_i are obtained by maximizing the log likelihood function (LLF)

$$\ell(\boldsymbol{P}, \boldsymbol{q}) = \log \theta_1 + \sum_{i=1}^{r} \sum_{j=1}^{c} n_{ij} \log p_{ij} + \sum_{i=1}^{r} m_i \log q_i$$

under (11). The MLE's under (11) are easily given by

$$\hat{p}_{ij} = \frac{n_{ij}}{n+m}, \qquad \hat{q}_i = \frac{m_i}{n+m}.$$

Now we consider the MLE's of p_{ij} and q_i under H_a. Let $q_i/p_{i\cdot} = \kappa$ $(\kappa > 0)$ for any i. Then, the condition (11) is written as $\sum\sum p_{ij} = (\kappa+1)^{-1}$. Under this condition, by maximizing the LLF

$$\ell^{(0)}(\boldsymbol{P}, \kappa) = \log \theta_1 + \sum_{i=1}^{r} \sum_{j=1}^{c} n_{ij} \log p_{ij} + \sum_{i=1}^{r} m_i \log p_{i\cdot} + m \log \kappa,$$

the MLE's of p_{ij}, κ and $q_i = \kappa p_{i\cdot}$ under H_a are obtained as

$$\hat{p}_{ij}^{(0)} = \frac{n n_{ij}(n_{i\cdot} + m_i)}{n_{i\cdot}(n+m)^2}, \qquad \hat{\kappa} = \frac{m}{n}, \qquad \hat{q}_i^{(0)} = \frac{m(n_{i\cdot} + m_i)}{(n+m)^2}. \tag{12}$$

Next we consider the asymptotic properties of these MLE's. In general, the p_{ij} and the q_i under H_a can be written as

$$p_{ij} = p_{ij}(\boldsymbol{\theta}), \qquad q_i = q_i(\boldsymbol{\theta}),$$

where t-dimensional parameter vector $\boldsymbol{\theta}$ belongs the parameter space Ω. Let $\hat{\boldsymbol{\theta}}$ be the MLE of $\boldsymbol{\theta}$, and let $\hat{p}_{ij} = p_{ij}(\hat{\boldsymbol{\theta}})$ and $\hat{q}_i = q_i(\hat{\boldsymbol{\theta}})$. To drive the asymptotic distribution of $\hat{p}_{ij}$ and $\hat{q}_i$ we need to assume the regularity conditions (see, e.g., Agresti [1]):

(i) $\boldsymbol{\theta}_0$ is not on the boundary of Ω,

(ii) all $p_{ij}(\boldsymbol{\theta}_0) > 0$, $q_i(\boldsymbol{\theta}_0) > 0$,

(iii) $p_{ij}(\boldsymbol{\theta})$ and $q_i(\boldsymbol{\theta})$ have continuous first-order partial derivatives in a neighborhood of $\boldsymbol{\theta}_0$, and

(iv) the Jacobian matrix $(\partial(p_{ij}, q_i)/\partial\boldsymbol{\theta})$ has full rank t at $\boldsymbol{\theta}_0$,

where $\boldsymbol{\theta}_0$ is the true value of $\boldsymbol{\theta}$. In the following problems, we assume (i) and (ii), but note that (iii) and (iv) are satisfied. We use the following vector and matrix notations.

$$\boldsymbol{N}^* = \mathrm{vec}(\tilde{\boldsymbol{N}}), \qquad \boldsymbol{P}^* = \mathrm{vec}(\tilde{\boldsymbol{P}}),$$

$$\boldsymbol{P}^{*1/2} = (\sqrt{p_{11}}, \ldots, \sqrt{p_{r1}}, \ldots, \sqrt{p_{1c}}, \ldots, \sqrt{p_{rc}}, \sqrt{q_1}, \ldots, \sqrt{q_r})',$$

$$\boldsymbol{D}_{p^*}^{1/2} = \mathrm{diag}(\boldsymbol{P}^{*1/2}).$$

Let $\boldsymbol{A}$ be $r(c+1) \times \{r(c+1)-1\}$ matrix such that the matrix $[\boldsymbol{A}|\boldsymbol{P}^{*1/2}]$ of order $r(c+1)$ is orthogonal, *i.e.*, it satisfies

$$\boldsymbol{A}\boldsymbol{A}' + \boldsymbol{P}^{*1/2}\boldsymbol{P}^{*1/2\prime} = \boldsymbol{I}, \qquad \boldsymbol{A}'\boldsymbol{A} = \boldsymbol{I}, \qquad \boldsymbol{A}'\boldsymbol{P}^{*1/2} = \boldsymbol{0}, \tag{13}$$

where $\boldsymbol{I}$ is the identity matrix. Also let

$$\boldsymbol{z} = (z_1, \ldots, z_{r(c+1)-1})' = \sqrt{n+m}\boldsymbol{A}'\boldsymbol{D}_{p^*}^{-1/2}\left(\frac{1}{n+m}\boldsymbol{N}^* - \boldsymbol{P}^*\right)$$

or

$$\frac{1}{n+m}\boldsymbol{N}^* = \boldsymbol{P}^* + \varepsilon \boldsymbol{D}_{p^*}^{1/2}\boldsymbol{A}\boldsymbol{z}, \tag{14}$$

where $\varepsilon = (n+m)^{-1/2}$. Then $\boldsymbol{z}$ has asymptotically multivariate normal with $E[\boldsymbol{z}] = \boldsymbol{0}$ and $Var[\boldsymbol{z}] = \boldsymbol{I}$ as $n+m \to \infty$ (see Rao [16, Chapter 6]). From (14), for $i = 1, \ldots, r$,

$$\frac{n_{ij}}{n+m} = p_{ij} + \varepsilon g_{ij} \qquad \text{for } j = 1, \ldots, c, \qquad \frac{m_i}{n+m} = q_i + \varepsilon h_i, \tag{15}$$

where

$$g_{ij} = \sqrt{p_{ij}} \sum_{k=1}^{r(c+1)-1} a_{(ij)k} z_k, \qquad h_i = \sqrt{q_i} \sum_{k=1}^{r(c+1)-1} a_{(i(c+1))k} z_k,$$

where $a_{(ij)k}$ stands for the $((ij), k)$-th element of $\boldsymbol{A}$ and the k-th column of $\boldsymbol{A}$ consists of $(a_{(11)k}, \ldots, a_{(r1)k}, a_{(12)k}, \ldots, a_{(r(c+1))k})'$. From (12) and (15), the perturbation expansions of $\hat{p}_{ij}^{(0)}$ and $\hat{q}_i^{(0)}$ are given by

$$\begin{aligned}\hat{p}_{ij}^{(0)} &= p_{ij} + \varepsilon\left\{-\frac{1}{\kappa+1} p_{ij} p_{i\cdot}^{-1}(\kappa g_{i\cdot} - h_i) + (\kappa+1) p_{ij} g_{\cdot\cdot} + g_{ij}\right\} \\ &\quad + \varepsilon^2\left[\frac{\kappa g_{i\cdot} - h_i}{(\kappa+1)p_{i\cdot}}\left\{\frac{p_{ij}}{p_{i\cdot}} g_{i\cdot} - (\kappa+1) p_{ij} g_{\cdot\cdot} - g_{ij}\right\} + (\kappa+1) g_{ij} g_{\cdot\cdot}\right] + O_p(\varepsilon^3),\end{aligned} \tag{16}$$

$$\hat{q}_i^{(0)} = q_i + \varepsilon\left\{\frac{\kappa}{\kappa+1}(g_{i\cdot} + h_i) + (\kappa+1) p_{i\cdot} h_{\cdot}\right\} + \varepsilon^2\{(g_{i\cdot} + h_i) h_{\cdot}\}, \tag{17}$$

respectively, where $(\kappa+1)^{-1} = \sum\sum p_{ij}$, $\kappa/(\kappa+1) = \sum q_i$, $g_{i\cdot} = \sum_j^c g_{ij}$, $g_{\cdot\cdot} = \sum\sum g_{ij}$ and $h_{\cdot} = \sum h_i$. By the asymptotic normality of z_k and the orthogonal conditions (13), the means and covariances of $\hat{p}_{ij}^{(0)}$ and $\hat{q}_i^{(0)}$ can be expressed as

$$E[\hat{p}_{ij}^{(0)}] = p_{ij} + O(\varepsilon^3), \qquad E[\hat{q}_i^{(0)}] = q_i, \tag{18}$$

$$Cov[\hat{p}_{ij}^{(0)}, \hat{p}_{i'j'}^{(0)}] = \varepsilon^2 \left\{ p_{ij}\delta_{(ij)(i'j')} - \frac{\kappa}{\kappa+1} p_{ij} p_{i'j'} p_{i\cdot}^{-1} \delta_{ii'} + (\kappa - 1) p_{ij} p_{i'j'} \right\}$$

$$+ O(\varepsilon^3), \tag{19}$$

$$Cov[\hat{q}_i^{(0)}, \hat{q}_{i'}^{(0)}] = \varepsilon^2 \left\{ \frac{\kappa}{\kappa+1} q_i \delta_{ii'} + \left(\frac{1}{\kappa} - 1 \right) q_i q_{i'} \right\} + O(\varepsilon^3), \tag{20}$$

$$Cov[\hat{p}_{ij}^{(0)}, \hat{q}_{i'}^{(0)}] = \varepsilon^2 \left\{ \frac{\kappa}{\kappa+1} p_{ij} \delta_{ii'} - 2 p_{ij} q_{i'} \right\} + O(\varepsilon^3), \tag{21}$$

$Var[\hat{p}_{ij}^{(0)}] = Cov[\hat{p}_{ij}^{(0)}, \hat{p}_{ij}^{(0)}]$ and $Var[\hat{q}_i^{(0)}] = Cov[\hat{q}_i^{(0)}, \hat{q}_i^{(0)}]$, where $\delta_{(ij)(i'j')}$ is 1 if $i = i'$ and $j = j'$, 0 if $i \neq i'$ or $j \neq j'$. We note that the error terms $O(\varepsilon^3)$ in (18)–(21) could be written as $O(\varepsilon^4)$, as in the usual statistical situation. The error terms given in the later sections also include the similar expression. Our results may be summarized as follows.

Theorem 4.1. *Let $\hat{p}_{ij}^{(0)}$ and $\hat{q}_i^{(0)}$ be the MLE's under the null hypothesis H_a. Then*

(i) *$\hat{p}_{ij}^{(0)}$ and $\hat{q}_i^{(0)}$ have the perturbation expansions given by* (16) *and* (17), *respectively,*

(ii) *$\sqrt{n+m}(\hat{p}_{ij}^{(0)} - p_{ij})$ and $\sqrt{n+m}(\hat{q}_i^{(0)} - q_i)$ are asymptotically distributed as* $N\left(0, p_{ij} - \frac{\kappa}{\kappa+1} p_{ij}^2 p_{i\cdot}^{-1} + (\kappa-1) p_{ij}^2\right)$ *and* $N\left(0, \frac{\kappa}{\kappa+1} q_i + \left(\frac{1}{\kappa} - 1\right) q_i^2\right)$, *respectively,*

(iii) *$E[\hat{p}_{ij}^{(0)}]$, $E[\hat{q}_i^{(0)}]$, $Cov[\hat{p}_{ij}^{(0)}, \hat{p}_{i'j'}^{(0)}]$, $Cov[\hat{q}_i^{(0)}, \hat{q}_{i'}^{(0)}]$ and $Cov[\hat{p}_{ij}^{(0)}, \hat{q}_{i'}^{(0)}]$ are expanded as* (18)–(21),

where $\kappa = \sum q_i / \sum \sum p_{ij}$.

Our purpose is to construct test statistics for the null hypothesis of (10). The log-likelihood ratio statistic T_{a1} can be defined as

$$T_{a1} = -2\{\ell(\hat{\boldsymbol{P}}, \hat{q}) - \ell^{(0)}(\hat{\boldsymbol{P}}^{(0)}, \hat{\kappa})\}$$

$$= 2 \sum_{i=1}^{r} \left\{ n_{i\cdot} \log \frac{n_{i\cdot}(n+m)}{n(n_{i\cdot} + m_i)} + m_i \log \frac{m_i(n+m)}{m(n_{i\cdot} + m_i)} \right\}. \tag{22}$$

Also the Wald statistic W_{a1} and the score statistic Q_{a1} can be written as

$$W_{a1} = \sum_{i=1}^{r} \frac{(n_{i\cdot} + m_i)(mn_{i\cdot} - nm_i)^2}{n_{i\cdot} m_i (n+m)^2}, \tag{23}$$

$$Q_{a1} = \sum_{i=1}^{r} \frac{(mn_{i\cdot} - nm_i)^2}{nm(n_{i\cdot} + m_i)}, \tag{24}$$

respectively.

Here we consider asymptotic properties of these statistics. Using (15), the perturbation expansions of T_{a1}, W_{a1} and Q_{a1} are derived. The first order terms of these statistics are equal and expressed as, for T_{a1},

$$T_{a1} = \left\{ \sum_{i=1}^{r} \frac{(\kappa g_{i\cdot} - h_i)^2}{\kappa(\kappa+1)p_{i\cdot}} - \frac{(\kappa+1)^2}{\kappa} g_{\cdot\cdot}^2 \right\} + O_p(\varepsilon), \tag{25}$$

where κ, $g_{i\cdot}$, $g_{\cdot\cdot}$ and h_i are the ones used in (16) and (17). Then, the first order term of T_{a1} can be represented as $\boldsymbol{z}'\boldsymbol{B}_{a1}\boldsymbol{z}$. We can see that $\boldsymbol{B}_{a1}^2 = \boldsymbol{B}_{a1}$ and $\mathrm{tr}(\boldsymbol{B}_{a1}) = (r-1)$. Further, we can check that

$$E[T_{a1}] = (r-1) + O(\varepsilon^2).$$

Our results may be summarized as follows.

Theorem 4.2. *Let T_{a1}, W_{a1} and Q_{a1} be the log-likelihood ratio, the Wald and the score statistics for the null hypothesis H_a, respectively. Then*
(i) *T_{a1}, W_{a1} and Q_{a1} are represented as* (22), (23) *and* (24), *respectively,*
(ii) *The perturbation expansions of T_{a1}, W_{a1} and Q_{a1} are all equal in the first order term, and they are given as the right-hand of* (25),
(iii) *The null distributions of T_{a1}, W_{a1} and Q_{a1} are asymptotically distributed as a χ^2-distribution χ^2_{r-1}.*

4.2 Model (2)

Suppose that the rows of $\tilde{\boldsymbol{N}}$, *i.e.*, $\boldsymbol{n}_i = (n_{i1}, \ldots, n_{ic}, m_i)'$ for $i = 1, \ldots, r$, have r independent multinomials with respective probability vectors $\boldsymbol{p}_i = (p_{i1}, \ldots, p_{ic}, q_i)'$ and trial size parameters $n_{i\cdot} + m_i$, where

$$\sum_{j=1}^{c} p_{ij} + q_i = 1 \qquad \text{for } i = 1, \ldots, r. \tag{26}$$

Then, the MLE's of p_{ij} and q_i under (26) are

$$\hat{p}_{ij} = \frac{n_{ij}}{n_{i\cdot} + m_i}, \qquad \hat{q}_i = \frac{m_i}{n_{i\cdot} + m_i}.$$

The joint density function of $\boldsymbol{n}_i$ $(i = 1, \ldots, r)$ and its LLF are written as

$$f_2(\boldsymbol{n}_1, \ldots, \boldsymbol{n}_r \,|\, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_r) = \theta_2 \prod_{i=1}^{r} \prod_{j=1}^{c} p_{ij}^{n_{ij}} \prod_{i=1}^{r} q_i^{m_i},$$

$$\ell(\boldsymbol{P}, \boldsymbol{q}) = \log \theta_2 + \sum_{i=1}^{r} \sum_{j=1}^{c} n_{ij} \log p_{ij} + \sum_{i=1}^{r} m_i \log q_i,$$

respectively, where $\theta_2 = \prod_{i=1}^{r} (n_{i\cdot} + m_i)!/(\prod_{i=1}^{r} \prod_{j=1}^{c} n_{ij}! \prod_{i=1}^{r} m_i!)$.

To derive the LR criterion for the null hypothesis H_a, we obtain the MLE's of p_{ij} and q_i under H_a. Let $q_i/p_{i\cdot} = \kappa$ $(\kappa > 0)$ for any i. Then, from (26), $p_{i\cdot} = (\kappa + 1)^{-1}$ and $q_i = \kappa/(\kappa + 1)$. By maximizing the LLF under the conditions $\sum_j p_{ij} = (\kappa + 1)^{-1}$ for $i = 1, \ldots, r$, the MLE's of the parameters under H_a are obtained as

$$\hat{p}_{ij}^{(0)} = \frac{n n_{ij}}{n_{i\cdot}(n + m)}, \qquad \hat{\kappa} = \frac{m}{n}, \qquad \hat{q}_i^{(0)} = \frac{m}{n + m}. \tag{27}$$

Now we consider the asymptotic properties of $\hat{p}_{ij}^{(0)}$ and $\hat{q}_i^{(0)}$. Assume the regularity conditions which are similar to them in the previous section. Let, for $i = 1, \ldots, r$,

$$\boldsymbol{p}_i^{1/2} = (\sqrt{p_{i1}}, \ldots, \sqrt{p_{ic}}, \sqrt{q_i})', \qquad \boldsymbol{D}_{p_i}^{1/2} = \mathrm{diag}(\boldsymbol{p}_i^{1/2}).$$

Similarly, let $\boldsymbol{A}^{(i)}$ be $(c + 1) \times c$ matrix such that the matrix $[\boldsymbol{A}^{(i)} | \boldsymbol{p}_i^{1/2}]$ of order $(c + 1)$ is orthogonal, *i.e.*, it satisfies

$$\boldsymbol{A}^{(i)} \boldsymbol{A}^{(i)\prime} + \boldsymbol{p}_i^{1/2} \boldsymbol{p}_i^{1/2\prime} = \boldsymbol{I}, \qquad \boldsymbol{A}^{(i)\prime} \boldsymbol{A}^{(i)} = \boldsymbol{I}, \qquad \boldsymbol{A}^{(i)\prime} \boldsymbol{p}_i^{1/2} = \boldsymbol{0}. \tag{28}$$

For $i = 1, \ldots, r$, consider

$$\frac{1}{n_{i\cdot} + m_i} \boldsymbol{n}_i = \boldsymbol{p}_i + \varepsilon_i \boldsymbol{D}_{p_i}^{1/2} \boldsymbol{A}^{(i)} \boldsymbol{z}^{(i)}, \tag{29}$$

where $\boldsymbol{z}^{(i)} = (z_1^{(i)}, \ldots, z_c^{(i)})'$ and $\varepsilon_i = (n_{i\cdot} + m_i)^{-1/2}$. Then $\boldsymbol{z}^{(i)}$ $(i = 1, \ldots, r)$ have asymptotically r independent multivariate normals with $E[\boldsymbol{z}^{(i)}] = \boldsymbol{0}$ and $Var[\boldsymbol{z}^{(i)}] = \boldsymbol{I}$ as $n_{i\cdot} + m_i \to \infty$ for $i = 1, \ldots, r$. Since $(n_{i\cdot} + m_i)$ are given for any i, $(n + m)$ is also given. Therefore, $\varepsilon = (n + m)^{-1/2}$ and $\rho_i = (n_{i\cdot} + m_i)/(n + m)$ are given constants, where $\sum \rho_i = 1$, $0 < \rho_i < 1$ for any i. Further, for $i = 1, \ldots, r$, since ε_i is denoted as $\varepsilon/\sqrt{\rho_i}$, from (29)

$$\frac{n_{ij}}{n + m} = \rho_i p_{ij} + \varepsilon g_{ij} \qquad \text{for } j = 1, \ldots, c, \qquad \frac{m_i}{n + m} = \rho_i q_i + \varepsilon h_i, \tag{30}$$

where

$$g_{ij} = \sqrt{\rho_i p_{ij}} \sum_{k=1}^{c} a_{jk}^{(i)} z_k^{(i)}, \qquad h_i = \sqrt{\rho_i q_i} \sum_{k=1}^{c} a_{(c+1)k}^{(i)} z_k^{(i)}.$$

Here $a_{jk}^{(i)}$ stands for the (j,k)-th element of $\boldsymbol{A}^{(i)}$. Using (27) and (30), the perturbation expansions of $\hat{p}_{ij}^{(0)}$ and $\hat{q}_i^{(0)}$ are given by

$$\hat{p}_{ij}^{(0)} = p_{ij} + \varepsilon \frac{1}{\rho_i p_{i\cdot}} \{p_{i\cdot} g_{ij} - p_{ij}(g_{i\cdot} - \rho_i g_{\cdot\cdot})\}$$

$$+ \varepsilon^2 \frac{1}{\rho_i^2 p_{i\cdot}^2}(g_{i\cdot} - \rho_i g_{\cdot\cdot})(p_{ij} g_{i\cdot} - p_{i\cdot} g_{ij}) + O_p(\varepsilon^3), \tag{31}$$

$$\hat{q}_i^{(0)} = q_i + \varepsilon h_{\cdot}, \tag{32}$$

respectively, where $g_{i\cdot} = \sum_j^c g_{ij}$, $g_{\cdot\cdot} = \sum_i^r g_{i\cdot}$ and $h_{\cdot} = \sum_i^r h_i$. By the asymptotic normality of $z_k^{(i)}$ and the orthogonal conditions (28), the means and covariances of $\hat{p}_{ij}^{(0)}$ and $\hat{q}_i^{(0)}$ can be expressed as

$$E[\hat{p}_{ij}^{(0)}] = p_{ij} + O(\varepsilon^3), \qquad E[\hat{q}_i^{(0)}] = q_i, \tag{33}$$

$$Cov[\hat{p}_{ij}^{(0)}, \hat{p}_{i'j'}^{(0)}] = \varepsilon^2 \left\{ \frac{p_{ij}(p_{i\cdot}\delta_{jj'} - p_{ij'})}{\rho_i p_{i\cdot}} \delta_{ii'} + p_{ij} p_{i'j'} \frac{1 - p_{i\cdot}}{p_{i\cdot}} \right\} + O(\varepsilon^3), \tag{34}$$

$$Cov[\hat{q}_i^{(0)}, \hat{q}_{i'}^{(0)}] = \varepsilon^2 \{q_i(1 - q_i)\}, \tag{35}$$

$$Cov[\hat{p}_{ij}^{(0)}, \hat{q}_{i'}^{(0)}] = \varepsilon^2 \{-p_{ij}(1 - p_{i\cdot})\} + O(\varepsilon^3), \tag{36}$$

$Var[\hat{p}_{ij}^{(0)}] = Cov[\hat{p}_{ij}^{(0)}, \hat{p}_{ij}^{(0)}]$ and $Var[\hat{q}_i^{(0)}] = Cov[\hat{q}_i^{(0)}, \hat{q}_i^{(0)}]$. It is seen that (35) and (36) are not depended on i'. Our results may be summarized as follows.

Theorem 4.3. *Let $\hat{p}_{ij}^{(0)}$ and $\hat{q}_i^{(0)}$ be the MLE's under the null hypothesis H_a. Then*

(i) $\hat{p}_{ij}^{(0)}$ *and* $\hat{q}_i^{(0)}$ *have the perturbation expansions given by* (31) *and* (32), *respectively,*

(ii) $\sqrt{n+m}(\hat{p}_{ij}^{(0)} - p_{ij})$ *and* $\sqrt{n+m}(\hat{q}_i^{(0)} - q_i)$ *are asymptotically distributed as* $N\left(0, \dfrac{p_{ij}(p_{i\cdot} - p_{ij} + \rho_i p_{ij} - \rho_i p_{ij} p_{i\cdot})}{\rho_i p_{i\cdot}}\right)$ *and* $N(0, q_i(1 - q_i))$, *respectively,*

(iii) $E[\hat{p}_{ij}^{(0)}]$, $E[\hat{q}_i^{(0)}]$, $Cov[\hat{p}_{ij}^{(0)}, \hat{p}_{i'j'}^{(0)}]$, $Cov[\hat{q}_i^{(0)}, \hat{q}_{i'}^{(0)}]$ *and* $Cov[\hat{p}_{ij}^{(0)}, \hat{q}_{i'}^{(0)}]$ *are expanded as* (33)–(36),

where $\rho_i = (n_{i\cdot} + m_i)/(n + m)$.

For the null hypothesis of (10), the log-likelihood ratio, the Wald and the score statistics (T_{a2}, W_{a2} and Q_{a2}) are defined as the same form with them in Model (1), *i.e.*, (22), (23) and (24). Using (30), the perturbation expansions of T_{a2}, W_{a2} and Q_{a2} are obtained. The first order terms of these statistics are equal and expressed as, for T_{a2},

$$T_{a2} = \frac{(\kappa + 1)^2}{\kappa} \left\{ \sum_{i=1}^r \frac{g_{i\cdot}^2}{\rho_i} - g_{\cdot\cdot}^2 \right\} + O_p(\varepsilon), \tag{37}$$

where $g_{i\cdot}$ is the one used in (31). Similarly, the first order term of T_{a2} can be represented as $\boldsymbol{z}'\boldsymbol{B}_{a2}\boldsymbol{z}$, where $\boldsymbol{B}_{a2}$ is satisfied $\boldsymbol{B}_{a2}^2 = \boldsymbol{B}_{a2}$ and $\mathrm{tr}(\boldsymbol{B}_{a2}) = (r-1)$. Further, we can check that

$$E[T_{a2}] = (r-1) + O(\varepsilon^2).$$

Our results may be summarized as follows.

THEOREM 4.4. *Let T_{a2}, W_{a2} and Q_{a2} be the log-likelihood ratio, the Wald and the score statistics for the null hypothesis H_a, respectively. Then*

(i) *T_{a2}, W_{a2} and Q_{a2} are represented as* (22), (23) *and* (24), *respectively,*

(ii) *The perturbation expansions of T_{a2}, W_{a2} and Q_{a2} are all equal in the first order term, and they are given as the right-hand of* (37),

(iii) *The null distributions of T_{a2}, W_{a2} and Q_{a2} are asymptotically distributed as χ^2_{r-1}.*

4.3 Model (3)

Suppose that the columns of $\tilde{\boldsymbol{N}}$, *i.e.*, $\boldsymbol{n}_j = (n_{1j}, \ldots, n_{rj})'$ for $j = 1, \ldots, c$ and $\boldsymbol{m} = (m_1, \ldots, m_r)'$, have $(c+1)$ independent multinomials with respective probability vectors $\boldsymbol{p}_j = (p_{1j}, \ldots, p_{rj})'$, $\boldsymbol{q} = (q_1, \ldots, q_r)'$ and trial size parameters $n_{\cdot j}$, m, where

$$\sum_{i=1}^{r} p_{ij} = 1 \qquad \text{for } j = 1, \ldots, c, \qquad \sum_{i=1}^{r} q_i = 1. \tag{38}$$

Note that the notations $\boldsymbol{n}_j$ and $\boldsymbol{p}_j$ are different from the ones used in the previous section, respectively. Then, the MLE's of p_{ij} and q_i under (38) are

$$\hat{p}_{ij} = \frac{n_{ij}}{n_{\cdot j}}, \qquad \hat{q}_i = \frac{m_i}{m}.$$

The joint density function of $\boldsymbol{n}_j$ $(j = 1, \ldots, c)$, $\boldsymbol{m}$ and its LLF are written as

$$f_3(\boldsymbol{n}_1, \ldots, \boldsymbol{n}_c, \boldsymbol{m} \,|\, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_r, \boldsymbol{q}) = \theta_3 \prod_{i=1}^{r} \prod_{j=1}^{c} p_{ij}^{n_{ij}} \prod_{i=1}^{r} q_i^{m_i},$$

$$\ell(\boldsymbol{P}, \boldsymbol{q}) = \log \theta_3 + \sum_{i=1}^{r} \sum_{j=1}^{c} n_{ij} \log p_{ij} + \sum_{i=1}^{r} m_i \log q_i,$$

respectively, where $\theta_3 = \prod_{j=1}^{c} n_{\cdot j}! m! / (\prod_{i=1}^{r} \prod_{j=1}^{c} n_{ij}! \prod_{i=1}^{r} m_i!)$.

To derive the LR criterion for the null hypothesis H_a, we consider the MLE's of p_{ij} and q_i under H_a. Let $q_i / p_{i\cdot} = \kappa$ $(\kappa > 0)$ for any i. Then, from (38), $\kappa = c^{-1}$. By maximizing the LLF

$$\ell^{(0)}(\boldsymbol{P}) = \log\theta_3 + \sum_i^r \sum_j^c n_{ij}\log p_{ij} + \sum_i^r m_i \log p_{i\cdot} - m\log c$$

under the conditions $\sum_i^r p_{ij} = 1$ for $j = 1, \ldots, c$, the MLE's of the parameters under H_a are obtained. However, we cannot explicitly obtain the MLE's, since they are the solutions of nonlinear equations

$$\frac{n_{ij}}{p_{ij}} + \frac{m_i}{p_{i\cdot}} - \lambda_j = 0 \qquad \text{for } i = 1, \ldots, r, \ \ j = 1, \ldots, c,$$

where λ_j is Lagrange multiplier and $\sum_j^c \lambda_j = n + m$. We denote them as $\hat{p}_{ij}^{(0)}$ and $\hat{q}_i^{(0)}$. For the null hypothesis of (10), the log likelihood ratio statistic T_{a3} can be written as

$$T_{a3} = 2\left(\sum_{i=1}^r \sum_{j=1}^c n_{ij}\log\frac{n_{ij}}{n_{\cdot j}\hat{p}_{ij}^{(0)}} + \sum_{i=1}^r m_i \log\frac{m_i}{m\hat{q}_i^{(0)}}\right). \tag{39}$$

Also the Wald statistic W_{a3} and the score statistic Q_{a3} can be written as

$$W_{a3} = -(n+m) + \sum_{i=1}^r \sum_{j=1}^c \frac{n_{\cdot j}^2 (\hat{p}_{ij}^{(0)})^2}{n_{ij}} + \sum_{i=1}^r \frac{m^2(\hat{q}_i^{(0)})^2}{m_i}, \tag{40}$$

$$Q_{a3} = \sum_{i=1}^r \frac{m_i}{\hat{q}_i^{(0)}}\left(\frac{m_i}{m} - \frac{1}{c}\sum_{j=1}^c \frac{n_{ij}}{n_{\cdot j}}\right), \tag{41}$$

respectively.

The asymptotic properties for the MLE of parameters under the null hypothesis H_a could be obtained by deriving linear approximations of the above non-linear equations. However these results have not obtained in simple forms, and so they are not given here.

5. Tests for Type (B)-redundancy of a column variable

In this section we consider the testing problem of deciding whether the $(c+1)$-th column variable is redundant for scaling of the column variable, for three models in §3. Using the statement (B4) in Theorem B, we may consider the following testing problem.

$$H_b : \frac{q_i}{q_\cdot} = \frac{p_{it}}{p_{\cdot t}} \ \ (i = 1, \ldots, r) \qquad \text{for some } t \in \{1, \ldots, c\}, \tag{42}$$

$$K_b : \text{not } H_b.$$

For Model (3), the null hypothesis can be expressed as

$$H_b : q_i = p_{it} \ \ (i = 1, \dots, r) \qquad \text{for some } t \in \{1, \dots, c\}.$$

Let $H_b^{(j)} : q_i/q_\cdot = p_{ij}/p_{\cdot j}$ for any i. Then, the null hypothesis H_b can be also expressed as follows.

$$H_b : \text{“}H_b^{(1)} \text{ or } \dots \text{ or } H_b^{(c)}\text{”}.$$

For the above testing problem, we propose a simultaneous test procedure. First, for each j, we construct a test statistic for $H_b^{(j)}$. Let $T^{(j)}$ be a test statistic for $H_b^{(j)}$. Secondly, we consider a test statistic $T_b = \min\{T^{(1)}, \dots, T^{(c)}\}$. This approach is similar to Dunnett's method in multiple comparison, which is a method to compare each new treatment with a control (see e.g., Hsu [9]). In order to get a critical point of T_b we need to develop an approximation method. For this, we will see that the statistics $T^{(j)}$ $(j = 1, \dots, c)$ have asymptotically chi-square distributions, respectively. It is expected that their asymptotic distribution will be a multivariate chi-square distribution (see e.g., Johnson and Kotz [10, Chapter 40]). To see this, for example, it needs to obtain an asymptotic expression for the joint characteristic function of $(T^{(1)}, \dots, T^{(c)})$. This problem and the asymptotic distribution of T_b are left as future works. However, in this section we will give asymptotic correlations of $T^{(j)}$ and $T^{(j')}$ for any j and j', which shall be fundamental for the approximation method.

5.1 Model (1)

We proceed using the notation given in §4.1. For fixed t, the MLE's of p_{it} and q_i under $H_b^{(t)}$ are obtained as

$$\hat{p}_{it}^{(0)} = \frac{n_{\cdot t}(n_{it} + m_i)}{(n+m)(n_{\cdot t} + m)}, \qquad \hat{q}_i^{(0)} = \frac{m(n_{it} + m_i)}{(n+m)(n_{\cdot t} + m)}.$$

The asymptotic properties of these MLE's can be given using (15). The perturbation expansions of $\hat{p}_{ij}^{(0)}$, $\hat{p}_{it}^{(0)}$ and $\hat{q}_i^{(0)}$ are given by

$$\hat{p}_{ij}^{(0)} = p_{ij} + \varepsilon g_{ij} \qquad (j \neq t),$$

$$\begin{aligned}\hat{p}_{it}^{(0)} &= p_{it} + \varepsilon \frac{1}{p_{\cdot t} + q_\cdot}\{q_i g_{\cdot t} - p_{it} h_\cdot + p_{\cdot t}(g_{it} + h_i)\} \\ &\quad + \varepsilon^2 \frac{(p_{it} h_\cdot - q_i g_{\cdot t})(g_{\cdot t} + h_\cdot) - (p_{\cdot t} h_\cdot - q_\cdot g_{\cdot t})(g_{it} + h_i)}{(p_{\cdot t} + q_\cdot)^2} + O_p(\varepsilon^3),\end{aligned}$$

$$\begin{aligned}\hat{q}_i^{(0)} &= q_i + \varepsilon \frac{1}{p_{\cdot t} + q_\cdot}\{p_{it} h_\cdot - q_i g_{\cdot t} + q_\cdot(g_{it} + h_i)\} \\ &\quad + \varepsilon^2 \frac{(p_{\cdot t} h_\cdot - q_\cdot g_{\cdot t})(g_{it} + h_i) - (p_{it} h_\cdot - q_i g_{\cdot t})(g_{\cdot t} + h_\cdot)}{(p_{\cdot t} + q_\cdot)^2} + O_p(\varepsilon^3),\end{aligned}$$

respectively, where g_{ij} and h_i are the ones given in (15) and $\varepsilon = (n+m)^{-1/2}$. For $j, j' \neq t$, the means and the covariances of $\hat{p}_{ij}^{(0)}$, $\hat{p}_{it}^{(0)}$ and $\hat{q}_i^{(0)}$ can be expressed as

$$E[\hat{p}_{ij}^{(0)}] = p_{ij}, \qquad E[\hat{p}_{it}^{(0)}] = p_{it} + O(\varepsilon^3), \qquad E[\hat{q}_i^{(0)}] = q_i + O(\varepsilon^3),$$

$$Cov[\hat{p}_{ij}^{(0)}, \hat{p}_{i'j'}^{(0)}] = \varepsilon^2 p_{ij}(\delta_{(ij)(i'j')} - p_{i'j'}),$$

$$Cov[\hat{p}_{ij}^{(0)}, \hat{p}_{i't}^{(0)}] = \varepsilon^2(-p_{ij}p_{i't}) + O(\varepsilon^3), \qquad Cov[\hat{p}_{ij}^{(0)}, \hat{q}_{i'}^{(0)}] = \varepsilon^2(-p_{ij}q_{i'}) + O(\varepsilon^3),$$

$$Cov[\hat{p}_{it}^{(0)}, \hat{p}_{i't}^{(0)}] = \varepsilon^2 \frac{p_{i't}}{p_{\cdot t} + q_\cdot}\{p_{\cdot t}(\delta_{ii'} - p_{it}) + q_i(1 - p_{\cdot t})\} + O(\varepsilon^3),$$

$$Cov[\hat{p}_{it}^{(0)}, \hat{q}_{i'}^{(0)}] = \varepsilon^2 \frac{p_{it}}{p_{\cdot t} + q_\cdot}\{q_\cdot(\delta_{ii'} - p_{i't}) - q_{i'}(1 + q_\cdot)\} + O(\varepsilon^3),$$

$$Cov[\hat{q}_i^{(0)}, \hat{q}_{i'}^{(0)}] = \varepsilon^2 \frac{q_{i'}}{p_{\cdot t} + q_\cdot}\{q_\cdot(\delta_{ii'} - q_i) + p_{it}(1 - q_\cdot)\} + O(\varepsilon^3),$$

respectively.

For the null hypothesis $H_b^{(t)}$ under the fixed t, the log-likelihood ratio statistic $T_{b1}^{(t)}$ can be written as

$$T_{b1}^{(t)} = 2\sum_{i=1}^{r}\left\{n_{it}\log\frac{n_{it}(n_{\cdot t} + m)}{n_{\cdot t}(n_{it} + m_i)} + m_i \log\frac{m_i(n_{\cdot t} + m)}{m(n_{it} + m_i)}\right\}. \tag{43}$$

Also the Wald statistic $W_{b1}^{(t)}$ and the score statistic $Q_{b1}^{(t)}$ can be written as

$$W_{b1}^{(t)} = \sum_{i=1}^{r}\frac{(n_{it} + m_i)(n_{it}m - n_{\cdot t}m_i)^2}{n_{it}m_i(n_{\cdot t} + m)^2}, \tag{44}$$

$$Q_{b1}^{(t)} = \sum_{i=1}^{r}\frac{(n_{it}m - n_{\cdot t}m_i)^2}{n_{\cdot t}m(n_{it} + m_i)}, \tag{45}$$

respectively. The perturbation expansions of $T_{b1}^{(t)}$, $W_{b1}^{(t)}$ and $Q_{b1}^{(t)}$ are given using (15). The first order terms of these statistics are equal, and they are expressed as, for $T_{b1}^{(t)}$,

$$T_{b1}^{(t)} = \sum_{i=1}^{r}\frac{(p_{it}h_i - q_i g_{it})^2}{p_{it}q_i(p_{it} + q_i)} - \frac{(p_{\cdot t}h_\cdot - q_\cdot g_{\cdot t})^2}{p_{\cdot t}q_\cdot(p_{\cdot t} + q_\cdot)} + O_p(\varepsilon). \tag{46}$$

Then the first order term of $T_{b1}^{(t)}$ can be represented as $\boldsymbol{z}'\boldsymbol{B}_{b1}\boldsymbol{z}$. We can see that $\boldsymbol{B}_{b1}^2 = \boldsymbol{B}_{b1}$ and $\mathrm{tr}(\boldsymbol{B}_{b1}) = (r-1)$. Further, we can check that

$$E[T_{b1}^{(t)}] = (r-1) + O(\varepsilon^2).$$

The covariance between $T_{b1}^{(t)}$ and $T_{b1}^{(t')}$ can be obtained as

$$Cov\left[T_{b1}^{(t)}, T_{b1}^{(t')}\right] = 2(r-1)\frac{(p_{\cdot t}+\delta_{tt'}q_{\cdot})(p_{\cdot t'}+\delta_{tt'}q_{\cdot})}{(p_{\cdot t}+q_{\cdot})(p_{\cdot t'}+q_{\cdot})} + O(\varepsilon).$$

5.2 Model (2)

We proceed using the notation given in §4.2. For fixed t, the MLE's of p_{it} and q_i under $H_b^{(t)}$ are obtained as

$$\hat{p}_{it}^{(0)} = \frac{n_{\cdot t}(n_{it}+m_i)}{(n_{i\cdot}+m_i)(n_{\cdot t}+m)}, \qquad \hat{q}_i^{(0)} = \frac{m(n_{it}+m_i)}{(n_{i\cdot}+m_i)(n_{\cdot t}+m)}.$$

The asymptotic properties of these MLE's can be given using (30). The perturbation expansions of $\hat{p}_{ij}^{(0)}$, $\hat{p}_{it}^{(0)}$ and $\hat{q}_i^{(0)}$ are given by

$$\hat{p}_{ij}^{(0)} = p_{ij} + \varepsilon\frac{1}{\rho_i}g_{ij} \qquad (j \neq t),$$

$$\hat{p}_{it}^{(0)} = p_{it} + \varepsilon\frac{\rho_i p_{it}(q_{\cdot}g_{\cdot t} - p_{\cdot t}h_{\cdot}) + p_{\cdot t}(\sum \rho_i p_{it})(g_{it}+h_i)}{\rho_i(p_{\cdot t}+q_{\cdot})\sum \rho_i p_{it}}$$
$$+ \varepsilon^2\frac{p_{\cdot t}(q_{\cdot}g_{\cdot t} - p_{\cdot t}h_{\cdot})\{(\sum \rho_i p_{it})(g_{it}+h_i) - \rho_i p_{it}(g_{\cdot t}+h_{\cdot})\}}{\rho_i(p_{\cdot t}+q_{\cdot})^2(\sum \rho_i p_{it})^2} + O_p(\varepsilon^3),$$

$$\hat{q}_i^{(0)} = q_i + \varepsilon\frac{q_{\cdot}(\sum \rho_i p_{it})(g_{it}+h_i) - \rho_i p_{it}(q_{\cdot}g_{\cdot t} - p_{\cdot t}h_{\cdot})}{\rho_i(p_{\cdot t}+q_{\cdot})\sum \rho_i p_{it}}$$
$$+ \varepsilon^2\frac{p_{\cdot t}(q_{\cdot}g_{\cdot t} - p_{\cdot t}h_{\cdot})\{\rho_i p_{it}(g_{\cdot t}+h_{\cdot}) - (\sum \rho_i p_{it})(g_{it}+h_i)\}}{\rho_i(p_{\cdot t}+q_{\cdot})^2(\sum \rho_i p_{it})^2} + O_p(\varepsilon^3),$$

respectively, where g_{ij} and h_i are the ones given in (30), $\rho_i = (n_{i\cdot}+m_i)/(n+m)$ and $\varepsilon = (n+m)^{-1/2}$. For $j, j' \neq t$, the means and covariances of $\hat{p}_{ij}^{(0)}$, $\hat{p}_{it}^{(0)}$ and $\hat{q}_i^{(0)}$ can be expressed as

$$E[\hat{p}_{ij}^{(0)}] = p_{ij}, \qquad E[\hat{p}_{it}^{(0)}] = p_{it} + O(\varepsilon^3), \qquad E[\hat{q}_i^{(0)}] = q_i + O(\varepsilon^3),$$

$$Cov[\hat{p}_{ij}^{(0)}, \hat{p}_{i'j'}^{(0)}] = \varepsilon^2\delta_{ii'}\frac{1}{\rho_i}p_{ij}(\delta_{jj'} - p_{i'j'}),$$

$$Cov[\hat{p}_{ij}^{(0)}, \hat{p}_{i't}^{(0)}] = \varepsilon^2\delta_{ii'}\left(-\frac{p_{ij}p_{it}}{\rho_i}\right) + O(\varepsilon^3),$$

$$Cov[\hat{p}_{ij}^{(0)}, \hat{q}_{i'}^{(0)}] = \varepsilon^2\delta_{ii'}\left(-\frac{p_{ij}q_i}{\rho_i}\right) + O(\varepsilon^3),$$

$$Cov[\hat{p}_{it}^{(0)}, \hat{p}_{i't}^{(0)}] = \varepsilon^2 \left\{ \frac{q_\cdot p_{it} p_{i't}}{(p_{\cdot t} + q_\cdot) \sum \rho_i p_{it}} + \delta_{ii'} \frac{p_{\cdot t} p_{it}(1 - p_{it} - q_i)}{\rho_i (p_{\cdot t} + q_\cdot)} \right\} + O(\varepsilon^3),$$

$$Cov[\hat{p}_{it}^{(0)}, \hat{q}_{i'}^{(0)}] = \varepsilon^2 \left\{ -\frac{q_\cdot p_{it} p_{i't}}{(p_{\cdot t} + q_\cdot) \sum \rho_i p_{it}} + \delta_{ii'} \frac{p_{\cdot t} q_i (1 - p_{it} - q_i)}{\rho_i (p_{\cdot t} + q_\cdot)} \right\} + O(\varepsilon^3),$$

$$Cov[\hat{q}_i^{(0)}, \hat{q}_{i'}^{(0)}] = \varepsilon^2 \left\{ \frac{q_\cdot p_{it} p_{i't}}{(p_{\cdot t} + q_\cdot) \sum \rho_i p_{it}} + \delta_{ii'} \frac{q_\cdot q_i (1 - p_{it} - q_i)}{\rho_i (p_{\cdot t} + q_\cdot)} \right\} + O(\varepsilon^3),$$

respectively.

For the null hypothesis $H_b^{(t)}$ under the fixed t, the log-likelihood ratio, the Wald and the score statistics ($T_{b2}^{(t)}$, $W_{b2}^{(t)}$ and $Q_{b2}^{(t)}$) can be written as the same form with them in Model (1), *i.e.*, (43), (44) and (45). The perturbation expansions of $T_{b2}^{(t)}$, $W_{b2}^{(t)}$ and $Q_{b2}^{(t)}$ are given using (30). The first order terms of these statistics are equal, and they are expressed as, for $T_{b2}^{(t)}$,

$$T_{b2}^{(t)} = \sum_{i=1}^{r} \frac{\{\rho_i p_{it}(q_\cdot g_{\cdot t} - p_{\cdot t} h_\cdot) - (\sum_i \rho_i p_{it})(q_\cdot g_{it} - p_{\cdot t} h_i)\}^2}{\rho_i p_{\cdot t} q_\cdot (p_{it} + q_i)(\sum_i \rho_i p_{it})^2} + O_p(\varepsilon). \qquad (47)$$

Similarly, the first order term of $T_{b2}^{(t)}$ can be represented as $\boldsymbol{z}'\boldsymbol{B}_{b2}\boldsymbol{z}$, where $\boldsymbol{B}_{b2}$ is satisfied $\boldsymbol{B}_{b2}^2 = \boldsymbol{B}_{b2}$ and $\mathrm{tr}(\boldsymbol{B}_{b2}) = (r-1)$. Further, we can check that

$$E[T_{b2}^{(t)}] = (r-1) + O(\varepsilon^2).$$

The covariance between $T_{b2}^{(t)}$ and $T_{b2}^{(t')}$ can be obtained as

$$Cov[T_{b2}^{(t)}, T_{b2}^{(t')}] = 2(r-1) \frac{(p_{\cdot t} + \delta_{tt'} q_\cdot)(p_{\cdot t'} + \delta_{tt'} q_\cdot)}{(p_{\cdot t} + q_\cdot)(p_{\cdot t'} + q_\cdot)} + O(\varepsilon).$$

5.3 Model (3)

We proceed using the notation given in §4.3. For fixed t, the MLE's of p_{it} and q_i under $H_b^{(t)}$ are obtained as

$$\hat{p}_{it}^{(0)} = \hat{q}_i^{(0)} = \frac{n_{it} + m_i}{n_{\cdot t} + m}.$$

Now, we consider the asymptotic properties of $\hat{p}_{ij}^{(0)}$, $\hat{p}_{it}^{(0)} (= \hat{q}_i^{(0)})$. Assume the regularity conditions which are similar to them in §4.1. Let, for $j = 1, \ldots, c$,

$$\boldsymbol{p}_j^{1/2} = (\sqrt{p_{1j}}, \ldots, \sqrt{p_{rj}})', \qquad \boldsymbol{D}_{p_j}^{1/2} = \mathrm{diag}(\boldsymbol{p}_j^{1/2}),$$
$$\boldsymbol{q}^{1/2} = (\sqrt{q_1}, \ldots, \sqrt{q_r})', \qquad \boldsymbol{D}_q^{1/2} = \mathrm{diag}(\boldsymbol{q}^{1/2}).$$

Similarly, let $\boldsymbol{A}^{(j)}$ and $\boldsymbol{A}^{(c+1)}$ be $r \times (r-1)$ matrices such that the matrices

$[\boldsymbol{A}^{(j)}|\boldsymbol{p}_j^{1/2}]$ and $[\boldsymbol{A}^{(c+1)}|\boldsymbol{q}^{1/2}]$ of order r are orthogonal. That is, it satisfies

$$\begin{aligned} \boldsymbol{A}^{(j)}\boldsymbol{A}^{(j)\prime} + \boldsymbol{p}_j^{1/2}\boldsymbol{p}_j^{1/2\prime} = \boldsymbol{I}, \qquad \boldsymbol{A}^{(j)\prime}\boldsymbol{A}^{(j)} = \boldsymbol{I}, \qquad \boldsymbol{A}^{(j)\prime}\boldsymbol{p}_j^{1/2} = \boldsymbol{0}, \\ \boldsymbol{A}^{(c+1)}\boldsymbol{A}^{(c+1)\prime} + \boldsymbol{q}^{1/2}\boldsymbol{q}^{1/2\prime} = \boldsymbol{I}, \qquad \boldsymbol{A}^{(c+1)\prime}\boldsymbol{A}^{(c+1)} = \boldsymbol{I}, \qquad \boldsymbol{A}^{(c+1)\prime}\boldsymbol{q}^{1/2} = \boldsymbol{0}. \end{aligned} \tag{48}$$

For $j = 1, \dots, c$, consider

$$\frac{1}{n_{\cdot j}}\boldsymbol{n}_j = \boldsymbol{p}_j + \varepsilon_j \boldsymbol{D}_{p_j}^{1/2}\boldsymbol{A}^{(j)}\boldsymbol{z}^{(j)}, \qquad \frac{1}{m}\boldsymbol{m} = \boldsymbol{q} + \varepsilon_{c+1}\boldsymbol{D}_q^{1/2}\boldsymbol{A}^{(c+1)}\boldsymbol{z}^{(c+1)}, \tag{49}$$

where $\boldsymbol{z}^{(j)} = (z_1^{(j)}, \dots, z_c^{(j)})'$, $\boldsymbol{z}^{(c+1)} = (z_1^{(c+1)}, \dots, z_c^{(c+1)})'$ and $\varepsilon_j = n_{\cdot j}^{-1/2}$ and $\varepsilon_{c+1} = m^{-1/2}$. Then $\boldsymbol{z}^{(j)}$ $(j = 1, \dots, c+1)$ have asymptotically $(c+1)$ independent multivariate normals with $E[\boldsymbol{z}^{(j)}] = \boldsymbol{0}$ and $Var[\boldsymbol{z}^{(j)}] = \boldsymbol{I}$ as $n_{\cdot j}, m \to \infty$ for $j = 1, \dots, c$. Since $n_{\cdot j}$ and m are given for any j, $(n+m)$ is also given. Therefore, $\varepsilon = (n+m)^{-1/2}$, $\rho_j = n_{\cdot j}/(n+m)$ and $\rho_{c+1} = m/(n+m)$ are given constants, where $\sum_{j=1}^{c+1}\rho_j = 1$, $0 < \rho_j < 1$ for any j. Further, since $\varepsilon_j = \varepsilon/\sqrt{\rho_j}$ for $j = 1, \dots, c+1$, from (49),

$$\frac{n_{ij}}{n+m} = \rho_j p_{ij} + \varepsilon g_{ij} \qquad \text{for } j = 1, \dots, c, \qquad \frac{m_i}{n+m} = \rho_{c+1} q_i + \varepsilon h_i, \tag{50}$$

where

$$g_{ij} = \sqrt{\rho_j p_{ij}} \sum_{k=1}^{r-1} a_{ik}^{(j)} z_k^{(j)}, \qquad h_i = \sqrt{\rho_{c+1} q_i} \sum_{k=1}^{r-1} a_{ik}^{(c+1)} z_k^{(c+1)}.$$

Here $a_{ik}^{(j)}$ stands for the (i,k)-th element of $\boldsymbol{A}^{(j)}$. Using (50), the perturbation expansions of $\hat{p}_{ij}^{(0)}$, $\hat{p}_{it}^{(0)} (= \hat{q}_i^{(0)})$ are given by

$$\hat{p}_{ij}^{(0)} = p_{ij} + \varepsilon \frac{1}{\sqrt{\rho_j}} g_{ij} \qquad (j \neq t), \qquad \hat{p}_{it}^{(0)} = p_{it} + \varepsilon \frac{1}{\rho_t + \rho_{c+1}} (g_{it} + h_i),$$

respectively. By the asymptotic normality of $z_k^{(j)}$ and the orthogonal conditions (48), for $j, j' \neq t$, the means and covariances of $\hat{p}_{ij}^{(0)}$ and $\hat{p}_{it}^{(0)}$ can be expressed as

$$E[\hat{p}_{ij}^{(0)}] = p_{ij}, \qquad E[\hat{p}_{it}^{(0)}] = p_{it}, \qquad Cov[\hat{p}_{ij}^{(0)}, \hat{p}_{i't}^{(0)}] = 0,$$

$$Cov[\hat{p}_{ij}^{(0)}, \hat{p}_{i'j'}^{(0)}] = \varepsilon^2 \delta_{jj'} \frac{1}{\rho_j} p_{ij}(\delta_{ii'} - p_{i'j'}),$$

$$Cov[\hat{p}_{it}^{(0)}, \hat{p}_{i't}^{(0)}] = \varepsilon^2 \frac{1}{\rho_t + \rho_{c+1}} p_{it}(\delta_{ii'} - p_{i't}),$$

respectively.

For the null hypothesis $H_b^{(t)}$ under the fixed t, the log-likelihood ratio, the Wald and the score statistics ($T_{b3}^{(t)}$, $W_{b3}^{(t)}$ and $Q_{b3}^{(t)}$) can be written as the same form with them in Model (1), *i.e.*, (43), (44) and (45). The perturbation expansions of $T_{b3}^{(t)}$, $W_{b3}^{(t)}$ and $Q_{b3}^{(t)}$ are given using (50). The first order terms of these statistics are equal, and they are expressed as, using the notation $T_{b3}^{(t)}$,

$$T_{b3}^{(t)} = \sum_{i=1}^{r} \frac{(\rho_{c+1} g_{it} - \rho_t h_i)^2}{\rho_t \rho_{c+1} (\rho_t + \rho_{c+1}) p_{it}} + O_p(\varepsilon). \tag{51}$$

Similarly, the first order term of $T_{b3}^{(t)}$ can be represented as $\boldsymbol{z}'\boldsymbol{B}_{b3}\boldsymbol{z}$, where $\boldsymbol{B}_{b3}$ is satisfied $\boldsymbol{B}_{b3}^2 = \boldsymbol{B}_{b3}$ and $\mathrm{tr}(\boldsymbol{B}_{b3}) = (r-1)$. Further, we can check that

$$E[T_{b3}^{(t)}] = (r-1) + O(\varepsilon^2).$$

The covariance between $T_{b3}^{(t)}$ and $T_{b3}^{(t')}$ can be obtained as

$$Cov[T_{b3}^{(t)}, T_{b3}^{(t')}] = 2(r-1)\frac{(\rho_t + \delta_{tt'}\rho_{c+1})(\rho_{t'} + \delta_{tt'}\rho_{c+1})}{(\rho_t + \rho_{c+1})(\rho_{t'} + \rho_{c+1})} + O(\varepsilon).$$

6. Examples and numerical experiments

In this section we give examples and numerical experiments that illustrate our results in §4.

6.1 Example

We consider using Table 9.7 in Greenacre [6]. The data is the number of doctorates classified by 12 fields from 1960 to 1976 in the USA. As the last column representing the year of 1976 is estimated frequencies, we removed this column from data table. Also, for convenience sake, replace row variables with column variables, *i.e.*, make an 8×12 contingency table that eight rows are variables denoting year taken a doctorate and twelve columns are fields of study. The 2-dimensional configuration by CA is illustrated in Figure 2. We may interpret that the 1-axis means the science field in the positive position and the literature field in the negative position, or the old year in the positive and the new year in the negative. Also, it is explained that the number of doctorates in the literature fields is few in the 1960's, but it is increasing in the 1970's. Further, it is seen that the number of doctorates in the engineering and mathematics fields are at its peak in 1971 and 1972, respectively. In this way, we can visually interpret the relations between years taken a doctorate, between fields of study and between year taken a doctorate and field of study.

Here, we consider the testing problem for Type (A)-redundancy of a column variable. By (A3) in Theorem A, if the coordinates of configuration of

Fig. 2. Optimal 2-dimensional configuration, by correspondence analysis, of the data of Table 9.7 in Greenacre [6].

Table 1. The values of three statistics and their p-value in Models (1) and (2) for testing problem of whether Biology is or not redundant for scaling of the row item: year.

	LR	Wald	Score
Value of statistic	4.577	4.529	4.602
p-value	0.711	0.717	0.708

a column variable in a low dimension are zero for any reduced dimension, it is redundant for scaling of the row variables. By Figure 2, it is seen that the coordinates of the 6-th column, *i.e.*, Biology is near zero. Therefore, we are interested in examining whether Biology is redundant for scaling of years taken a doctorate. As the test statistics in Models (1) and (2) are denoted as the same form, the values of them are simultaneously obtained and assessed by the asymptotically χ^2_7. These values are tabulated in Table 1. Since all p-values are about 0.71, the null hypothesis of (10) may be accepted. That is, we cannot say that Biology is not redundant for scaling of year taken a doctorate.

Table 2. The values of three statistics and their p-value in Models (1) and (2) for testing problem of whether Economics is or not redundant for scaling of the row item: year.

	LR	Wald	Score
Value of statistic	24.96	24.62	25.21
p-value	0.001	0.001	0.001

Table 3. The values of mean and percentiles of several n in Model (1) for test statistic of whether a column variable is or not redundant for scaling of the row variables.

	χ^2	LR			Wald			Score		
n		50	100	500	50	100	500	50	100	500
Mean	2	2.11	2.23	2.17	2.70	2.72	2.22	2.06	2.17	2.16
95%	5.99	6.26	6.63	6.76	9.57	8.81	6.98	5.90	6.14	6.73
99%	9.21	8.61	10.22	9.30	14.36	19.92	10.14	8.31	9.60	9.08

Similarly, we consider if 10-th column, *i.e.*, Economics is redundant for scaling of years taken a doctorate. The values of test statistics are tabulated in Table 2. Since all p-values are about 0.001, the null hypothesis of (10) may be rejected. That is, we have a conclusion that Economics is not redundant for scaling of year taken a doctorate.

6.2 Numerical experiments

In Model (1), we consider a numerical experiment on a 3×4 contingency table with the following artificial probability matrix satisfying the null hypothesis of (10) for a multinomial distribution

$$\boldsymbol{P} = \begin{pmatrix} 0.0833 & 0.0556 & 0.1111 & 0.0833 \\ 0.0278 & 0.0833 & 0.0556 & 0.0556 \\ 0.1666 & 0.1111 & 0.0556 & 0.1111 \end{pmatrix}.$$

Let n be the sample size. Our interests are to examine asymptotic performances of various formulas under Model (1) in §4 through the numerical experiment for different values of n. In Table 3, the simulated values of mean and percentiles are tabulated for several values of n. The values in the *LR*, *Wald* and *Score* columns are obtained from 1000 iterations with multinomial parameter matrix $\boldsymbol{P}$. The values in the χ^2 column are mean and percentiles of the exact χ^2_2. We observe that mean and percentiles of the LR and the score

Table 4. The values of mean and percentiles of several n in Model (2) for test statistic of whether a column variable is or not redundant for scaling of the row variables.

	χ^2	LR			Wald			Score		
n		50	100	500	50	100	500	50	100	500
Mean	2	1.81	2.09	2.02	2.30	2.37	2.06	1.76	2.05	2.01
95%	5.99	5.02	6.01	6.00	6.97	7.45	6.02	4.80	5.95	5.95
99%	9.21	7.51	9.41	8.38	13.85	12.17	9.17	7.00	8.75	8.24

statistics are considerably accurate for any n. However, for small n, such accuracy cannot be observed for them of Wald statistic.

In Model (2), we consider a numerical experiment on a 3×4 contingency table, which has the following artificial probability vectors satisfying the null hypothesis of (42) for three independent multinomial distributions

$$\boldsymbol{p}_1 = \begin{pmatrix} 0.250 \\ 0.167 \\ 0.333 \\ 0.250 \end{pmatrix}, \qquad \boldsymbol{p}_2 = \begin{pmatrix} 0.125 \\ 0.375 \\ 0.250 \\ 0.250 \end{pmatrix}, \qquad \boldsymbol{p}_3 = \begin{pmatrix} 0.375 \\ 0.250 \\ 0.125 \\ 0.250 \end{pmatrix}.$$

Let n_i $(i = 1, 2, 3)$ be each sample size and n the total sample size. In Table 4, the simulated values of mean and percentiles are tabulated for several values of n. It is noted that these values are independent of the selection of n_i. The degree of freedom in this case is 2. We observe that mean and percentiles of the LR and the score statistics are considerably accurate for $n = 100,\ 500$. However, for $n = 50$, such accuracy cannot be observed. Similarly, for small n, such accuracy cannot be observed for them of the Wald statistic.

7. Proofs of theorems

This section gives proofs of theorems stated in § 3. We use the following matrix notations.

$$\boldsymbol{P} = \left(\frac{p_{ij}}{p_{\cdot\cdot}}\right) : r \times c, \qquad \tilde{\boldsymbol{P}} = \left(\frac{\tilde{p}_{ij}}{p_{\cdot\cdot} + q_{\cdot}}\right) : r \times (c+1),$$

$$\boldsymbol{\Delta}_r = \mathrm{diag}\left(\frac{p_{i\cdot}}{p_{\cdot\cdot}}\right), \qquad \boldsymbol{\Delta}_c = \mathrm{diag}\left(\frac{p_{\cdot j}}{p_{\cdot\cdot}}\right),$$

$$\tilde{\boldsymbol{\Delta}}_r = \mathrm{diag}\left(\frac{p_{i\cdot} + q_i}{p_{\cdot\cdot} + q_{\cdot}}\right), \qquad \tilde{\boldsymbol{\Delta}}_c = \left[\begin{array}{c|c} \dfrac{p_{\cdot\cdot}}{p_{\cdot\cdot} + q_{\cdot}} \boldsymbol{\Delta}_c & \boldsymbol{0} \\ \hline \boldsymbol{0}' & \dfrac{q_{\cdot}}{p_{\cdot\cdot} + q_{\cdot}} \end{array}\right],$$

where $\tilde{p}_{ij} = p_{ij}$ for $j = 1, \ldots, c$ and $\tilde{p}_{i(c+1)} = q_i$. Note that each element in $\boldsymbol{P}$ and $\tilde{\boldsymbol{P}}$ is divided by the sums of all the elements. Let $\boldsymbol{\Theta}$ and $\tilde{\boldsymbol{\Theta}}$ be the matrices in the equation (7), respectively. In CA, the eigenvalue-eigenvector problem of $\tilde{\boldsymbol{\Theta}}$ can be also expressed as that of $\tilde{\boldsymbol{\Delta}}_r^{-1}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{\Delta}}_c^{-1}\tilde{\boldsymbol{P}}'$. Further, the non-zero eigenvalue $\tilde{\lambda}_\alpha$ of $\tilde{\boldsymbol{\Theta}}$ is the same as that of $\tilde{\boldsymbol{\Delta}}_r^{-1}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{\Delta}}_c^{-1}\tilde{\boldsymbol{P}}'$. Let $\tilde{\boldsymbol{v}}_\alpha$ be an eigenvector of $\tilde{\boldsymbol{\Delta}}_r^{-1}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{\Delta}}_c^{-1}\tilde{\boldsymbol{P}}'$ with $\tilde{\boldsymbol{v}}_\alpha'\tilde{\boldsymbol{\Delta}}_r\tilde{\boldsymbol{v}}_\beta = \delta_{\alpha\beta}$. Then the eigenvector $\tilde{\boldsymbol{v}}_\alpha$ has the following relations for $\alpha = 1, \ldots, \tilde{d}$.

$$\sqrt{\tilde{\lambda}_\alpha}\tilde{\boldsymbol{v}}_\alpha = \tilde{\boldsymbol{\Delta}}_r^{-1}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{\mu}}_\alpha, \tag{52}$$

$$\boldsymbol{1}_r'\tilde{\boldsymbol{\Delta}}_r\tilde{\boldsymbol{v}}_\alpha = 0. \tag{53}$$

7.1 Proof of Theorem A

Proof of Theorem A. From (8) it is obvious that (A3) $\Leftrightarrow$ (A4). The matrix $\tilde{\boldsymbol{\Theta}}$ may be written as

$$\tilde{\boldsymbol{\Theta}} = \left[\begin{array}{c|c} \dfrac{p_{\cdot\cdot}}{p_{\cdot\cdot}+q_\cdot}\boldsymbol{\Delta}_c^{-1}\boldsymbol{P}'\tilde{\boldsymbol{\Delta}}_r^{-1}\boldsymbol{P} & \dfrac{q_\cdot}{p_{\cdot\cdot}+q_\cdot}\boldsymbol{\Delta}_c^{-1}\boldsymbol{P}'\tilde{\boldsymbol{\Delta}}_r^{-1}\boldsymbol{q} \\ \hline \dfrac{p_{\cdot\cdot}}{p_{\cdot\cdot}+q_\cdot}\boldsymbol{q}'\tilde{\boldsymbol{\Delta}}_r^{-1}\boldsymbol{P} & \dfrac{q_\cdot}{p_{\cdot\cdot}+q_\cdot}\boldsymbol{q}'\tilde{\boldsymbol{\Delta}}_r^{-1}\boldsymbol{q} \end{array}\right],$$

where $\boldsymbol{q} = (q_i/q_\cdot, \ldots, q_r/q_\cdot)'$. In the following we prove in order of (A4) $\Leftrightarrow$ (A5), (A2) $\Leftrightarrow$ (A4) and (A1) $\Leftrightarrow$ (A5). Then, all equivalent relations are shown.

First we prove (A4) $\Rightarrow$ (A5). Considering the $(c+1)$-th components of $\tilde{\boldsymbol{\Theta}}\tilde{\boldsymbol{\mu}}_\alpha = \tilde{\lambda}_\alpha\tilde{\boldsymbol{\mu}}_\alpha$, we have

$$\frac{p_{\cdot\cdot}}{p_{\cdot\cdot}+q_\cdot}\boldsymbol{q}'\tilde{\boldsymbol{\Delta}}_r^{-1}\boldsymbol{P}\tilde{\boldsymbol{\mu}}_{1\alpha} = 0, \tag{54}$$

where $\tilde{\boldsymbol{\mu}}_\alpha = (\tilde{\boldsymbol{\mu}}_{1\alpha}', \tilde{\mu}_{(c+1)\alpha})'$. Further, by (52) and (A4),

$$\sqrt{\tilde{\lambda}_\alpha}\tilde{\boldsymbol{v}}_\alpha = \frac{p_{\cdot\cdot}}{p_{\cdot\cdot}+q_\cdot}\tilde{\boldsymbol{\Delta}}_r^{-1}\boldsymbol{P}\tilde{\boldsymbol{\mu}}_{1\alpha}. \tag{55}$$

These imply that for $\boldsymbol{q}'\tilde{\boldsymbol{v}}_\alpha = 0$, *i.e.*

$$\boldsymbol{1}_r'\boldsymbol{\Delta}_q\tilde{\boldsymbol{v}}_\alpha = 0, \tag{56}$$

where $\boldsymbol{\Delta}_q = \mathrm{diag}(q_i/q_\cdot)$. On the other hand, from the definition of $\tilde{\boldsymbol{v}}_\alpha$, we have $\boldsymbol{1}'\tilde{\boldsymbol{\Delta}}_r\tilde{\boldsymbol{v}}_\alpha = 0$ $(\alpha = 1, \ldots, \tilde{d})$. Therefore,

$$\boldsymbol{1}_r'(\boldsymbol{\Delta}_q\tilde{\boldsymbol{\Delta}}_r^{-1} - \boldsymbol{I})\tilde{\boldsymbol{\Delta}}_r\tilde{\boldsymbol{v}}_\alpha = 0 \qquad \text{for } \alpha = 1, \ldots, \tilde{d}.$$

This expression can be written as

$$\boldsymbol{g}'\tilde{\boldsymbol{\Delta}}_r\tilde{\boldsymbol{v}}_\alpha = 0 \qquad \text{for } \alpha = 1, \ldots, \tilde{d},$$

where $\boldsymbol{g} = (\tilde{\varDelta}_r^{-1}\varDelta_q - \boldsymbol{I})\boldsymbol{1}_r$. This implies that $\boldsymbol{g} = \boldsymbol{0}$ or $\kappa\boldsymbol{1}_r$ (κ is a constant). If $\boldsymbol{g} = \kappa\boldsymbol{1}$, we get $0 = \boldsymbol{1}_r'\tilde{\varDelta}_r\boldsymbol{g} = \kappa$, which gives a contradiction. Therefore, $\boldsymbol{g} = 0$, and hence $(\varDelta_q - \tilde{\varDelta}_r)\boldsymbol{1}_r = \boldsymbol{0}$, *i.e.*, (A5).

Next we prove (A5) $\Rightarrow$ (A4). By (A5), $\tilde{\boldsymbol{\Theta}}$ is written as

$$\frac{1}{\kappa+1}\left[\begin{array}{c|c} \varDelta_c^{-1}\boldsymbol{P}'\varDelta_r^{-1}\boldsymbol{P} & \kappa\boldsymbol{1}_c \\ \hline \boldsymbol{p}_c' & \kappa \end{array}\right],$$

where $\boldsymbol{p}_c = \varDelta_c\boldsymbol{1}_c$. Then

$$\frac{1}{\kappa+1}\left[\begin{array}{c} \varDelta_c^{-1}\boldsymbol{P}'\varDelta_r^{-1}\boldsymbol{P}\tilde{\boldsymbol{\mu}}_{1\alpha} + \kappa\boldsymbol{1}_c\tilde{\mu}_{(c+1)\alpha} \\ \hline \boldsymbol{p}_c'\tilde{\boldsymbol{\mu}}_{1\alpha} + \kappa\tilde{\mu}_{(c+1)\alpha} \end{array}\right] = \tilde{\lambda}_\alpha\left[\begin{array}{c} \tilde{\boldsymbol{\mu}}_{1\alpha} \\ \hline \tilde{\mu}_{(c+1)\alpha} \end{array}\right]. \tag{57}$$

Also, from the restriction $\boldsymbol{1}_{c+1}'\tilde{\varDelta}_c\tilde{\boldsymbol{\mu}}_\alpha = 0$, $\boldsymbol{1}_c'\varDelta_c\tilde{\boldsymbol{\mu}}_{1\alpha} + \kappa\tilde{\mu}_{(c+1)\alpha} = 0$, *i.e.*, $\boldsymbol{p}_c'\tilde{\boldsymbol{\mu}}_{1\alpha} + \kappa\tilde{\mu}_{(c+1)\alpha} = 0$. Therefore, by (57), $\tilde{\mu}_{(c+1)\alpha} = 0$ for α where $\tilde{\lambda}_\alpha \neq 0$, and hence (A5) $\Rightarrow$ (A4).

Note: By the above result, the first c component vector in (57) can be determined as $\frac{1}{\kappa+1}\varDelta_c^{-1}\boldsymbol{P}'\varDelta_r^{-1}\boldsymbol{P}\tilde{\boldsymbol{\mu}}_{1\alpha} = \tilde{\lambda}_\alpha\tilde{\boldsymbol{\mu}}_{1\alpha}$. This equation means that $\tilde{\boldsymbol{\mu}}_{1\alpha}$ is an eigenvector for an eigenvalue $(\kappa+1)\tilde{\lambda}_\alpha$ of the matrix $\varDelta_c^{-1}\boldsymbol{P}'\varDelta_r^{-1}\boldsymbol{P}$, *i.e.*, $\tilde{\boldsymbol{\mu}}_{1\alpha} = \boldsymbol{\mu}_\alpha$. Further, since (A5) means that an additional column vector is denoted as a linear combination of the other column vectors, it holds that the rank of $\tilde{\varDelta}_c^{-1}\tilde{\boldsymbol{P}}'\tilde{\varDelta}_r^{-1}\tilde{\boldsymbol{P}}$ is equal to the one of $\varDelta_c^{-1}\boldsymbol{P}'\varDelta_r^{-1}\boldsymbol{P}$. These properties hold under each of the conditions (A1)–(A5) which are shown to be equivalent each other, later.

We prove (A4) $\Rightarrow$ (A2). By (A4), the coordinate $\tilde{\xi}_{i\alpha}$ in (9) is expressed as

$$\tilde{\xi}_{i\alpha} = \frac{1}{p_{i\cdot} + q_i}\sum_{j=1}^{c} p_{ij}\tilde{\mu}_{j\alpha}.$$

Further, since (A4) $\Leftrightarrow$ (A5) and $\tilde{\boldsymbol{\mu}}_{1\alpha} = \boldsymbol{\mu}_\alpha$, for any i and α ($\tilde{\lambda}_\alpha \neq 0$),

$$\tilde{\xi}_{i\alpha} = \frac{1}{(\kappa+1)p_{i\cdot}}\sum_{j=1}^{c} p_{ij}\mu_{j\alpha} = \frac{1}{\kappa+1}\xi_{i\alpha}.$$

Therefore, (A4) $\Rightarrow$ (A2).

Next we prove (A2) $\Rightarrow$ (A4). The condition (A2) can be expressed in the terms of the vectors $\tilde{\boldsymbol{\mu}}_\alpha$ and $\boldsymbol{\mu}_\alpha$, as follows.

$$\tilde{\varDelta}_r^{-1}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{\mu}}_\alpha = \frac{1}{\kappa+1}\varDelta_r^{-1}\boldsymbol{P}\boldsymbol{\mu}_\alpha. \tag{58}$$

Further, the equation (7) is expressed as

$$\Delta_c^{-1}\boldsymbol{P}'\Delta_r^{-1}\boldsymbol{P}\boldsymbol{\mu}_\alpha = \lambda_\alpha\boldsymbol{\mu}_\alpha, \tag{59}$$

$$\tilde{\Delta}_c^{-1}\tilde{\boldsymbol{P}}'\tilde{\Delta}_r^{-1}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{\mu}}_\alpha = \tilde{\lambda}_\alpha\tilde{\boldsymbol{\mu}}_\alpha. \tag{60}$$

Then, from (60),

$$\Delta_c^{-1}\boldsymbol{P}'\tilde{\Delta}_r^{-1}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{\mu}}_\alpha = \tilde{\lambda}_\alpha\tilde{\boldsymbol{\mu}}_{1\alpha} \tag{61}$$

and by (58) and (59),

$$\Delta_c^{-1}\boldsymbol{P}'\tilde{\Delta}_r^{-1}\tilde{\boldsymbol{P}}\tilde{\boldsymbol{\mu}}_\alpha = \frac{\lambda_\alpha}{\kappa+1}\boldsymbol{\mu}_\alpha. \tag{62}$$

The relations (61) and (62) imply that $\tilde{\boldsymbol{\mu}}_{1\alpha}$ is proportional to $\boldsymbol{\mu}_\alpha$. Also, noting $\mathbf{1}'\Delta_c\boldsymbol{\mu}_\alpha = 0$ and $\mathbf{1}'\tilde{\Delta}_c\tilde{\boldsymbol{\mu}}_\alpha = 0$, we obtain

$$\mathbf{1}'\tilde{\Delta}_c\tilde{\boldsymbol{\mu}}_\alpha = \frac{p_{\cdot\cdot}}{p_{\cdot\cdot}+q_\cdot}\mathbf{1}'\Delta_c\tilde{\boldsymbol{\mu}}_{1\alpha} + \frac{q_\cdot}{p_{\cdot\cdot}+q_\cdot}\tilde{\mu}_{(c+1)\alpha} = 0.$$

As $\tilde{\boldsymbol{\mu}}_{1\alpha}$ is proportional to $\boldsymbol{\mu}_\alpha$, $\mathbf{1}'\Delta_c\tilde{\boldsymbol{\mu}}_{1\alpha} = 0$. Therefore, $\tilde{\mu}_{(c+1)\alpha} = 0$ and (A2) $\Rightarrow$ (A4).

We prove (A5) $\Rightarrow$ (A1). By (A5), the chi-square distance $\tilde{\tau}_r^2(i,i')$ between the i-th and the i'-th row variables based on the profile matrix $\tilde{\boldsymbol{P}}$ is expressed as

$$\begin{aligned}
\tilde{\tau}_r^2(i,i') &= \sum_j^c \frac{(\kappa+1)p_{\cdot\cdot}}{p_{\cdot j}}\left\{\frac{p_{ij}}{(\kappa+1)p_{i\cdot}} - \frac{p_{i'j}}{(\kappa+1)p_{i'\cdot}}\right\}^2 \\
&\quad + \frac{(\kappa+1)p_{\cdot\cdot}}{\kappa p_{\cdot\cdot}}\left\{\frac{\kappa p_{i\cdot}}{(\kappa+1)p_{i\cdot}} - \frac{\kappa p_{i'\cdot}}{(\kappa+1)p_{i'\cdot}}\right\}^2 \\
&= \frac{1}{\kappa+1}\sum_j^c \frac{p_{\cdot\cdot}}{p_{\cdot j}}\left(\frac{p_{ij}}{p_{i\cdot}} - \frac{p_{i'j}}{p_{i'\cdot}}\right)^2 \\
&= \frac{1}{\kappa+1}\tau_r^2(i,i').
\end{aligned}$$

As this relation holds for any i and i', we get (A5) $\Rightarrow$ (A1).

Finally we prove (A1) $\Rightarrow$ (A5). Let $\boldsymbol{e}_{ii'}$ be the $r \times 1$ vector which the i-th element is 1, the i'-th element -1 and the others 0. Then the chi-square distances between the i-th row variable and the i'-th row variable are expressed as

$$\tau_r^2(i,i') = \boldsymbol{e}_{ii'}'\boldsymbol{\Gamma}\boldsymbol{e}_{ii'}, \qquad \tilde{\tau}_r^2(i,i') = \boldsymbol{e}_{ii'}'\tilde{\boldsymbol{\Gamma}}\boldsymbol{e}_{ii'},$$

where $\boldsymbol{\Gamma} = \Delta_r^{-1}\boldsymbol{P}\Delta_c^{-1}\boldsymbol{P}'\Delta_r^{-1}$ and $\tilde{\boldsymbol{\Gamma}} = \tilde{\Delta}_r^{-1}\tilde{\boldsymbol{P}}\tilde{\Delta}_c^{-1}\tilde{\boldsymbol{P}}'\tilde{\Delta}_r^{-1}$. The (i,i')-th elements of $\boldsymbol{\Gamma}$ and $\tilde{\boldsymbol{\Gamma}}$ are written as

$$\gamma_{ii'} = \frac{p_{\cdot\cdot}}{p_{i\cdot}p_{i'\cdot}}\sum_{j=1}^{c}\frac{p_{ij}p_{i'j}}{p_{\cdot j}}, \qquad \tilde{\gamma}_{ii'} = \frac{p_{\cdot\cdot}+q_{\cdot}}{(p_{i\cdot}+q_i)(p_{i'\cdot}+q_{i'})}\left(\sum_{j=1}^{c}\frac{p_{ij}p_{i'j}}{p_{\cdot j}}+\frac{q_iq_{i'}}{q_{\cdot}}\right),$$

respectively. These elements can be related as

$$\begin{aligned}\tilde{\gamma}_{ii'} &= \frac{p_{\cdot\cdot}+q_{\cdot}}{p_{i\cdot}p_{i'\cdot}}\frac{p_{i\cdot}}{p_{i\cdot}+q_i}\frac{p_{i'\cdot}}{p_{i'\cdot}+q_{i'}}\left(\sum_{j=1}^{c}\frac{p_{ij}p_{i'j}}{p_{\cdot j}}+\frac{q_iq_{i'}}{q_{\cdot}}\right)\\ &= \frac{(p_{\cdot\cdot}+q_{\cdot})p_{i\cdot}p_{i'\cdot}}{(p_{i\cdot}+q_i)(p_{i'\cdot}+q_{i'})}\gamma_{ii'}+\frac{(p_{\cdot\cdot}+q_{\cdot})q_iq_{i'}}{q_{\cdot}(p_{i\cdot}+q_i)(p_{i'\cdot}+q_{i'})}.\end{aligned}$$

As the matrices $\boldsymbol{\Gamma}$ and $\tilde{\boldsymbol{\Gamma}}$ are symmetric,

$$\tau_r^2(i,i') = \gamma_{ii} - 2\gamma_{ii'} + \gamma_{i'i'}, \qquad \tilde{\tau}_r^2(i,i') = \tilde{\gamma}_{ii} - 2\tilde{\gamma}_{ii'} + \tilde{\gamma}_{i'i'}.$$

Therefore,

$$\begin{aligned}\tilde{\tau}_r^2(i,i') &= \frac{p_{\cdot\cdot}+q_{\cdot}}{p_{\cdot\cdot}}\left\{\frac{p_{i\cdot}^2\gamma_{ii}}{(p_{i\cdot}+q_i)^2}-\frac{2p_{i\cdot}p_{i'\cdot}\gamma_{ii'}}{(p_{i\cdot}+q_i)(p_{i'\cdot}+q_{i'})}+\frac{p_{i'\cdot}^2\gamma_{i'i'}}{(p_{i'\cdot}+q_{i'})^2}\right\}\\ &\quad+\frac{p_{\cdot\cdot}+q_{\cdot}}{q_{\cdot}}\left(\frac{q_i}{p_{i\cdot}+q_i}-\frac{q_{i'}}{p_{i'\cdot}+q_{i'}}\right)^2.\end{aligned}$$

Since $\tilde{\tau}_r^2(i,i') = \dfrac{1}{\kappa+1}\tau_r^2(i,i')$,

$$\frac{1}{\kappa+1} = \frac{(p_{\cdot\cdot}+q_{\cdot})p_{i\cdot}^2}{p_{\cdot\cdot}(p_{i\cdot}+q_i)^2} = \frac{(p_{\cdot\cdot}+q_{\cdot})p_{i\cdot}p_{i'\cdot}}{p_{\cdot\cdot}(p_{i\cdot}+q_i)(p_{i'\cdot}+q_{i'})} = \frac{(p_{\cdot\cdot}+q_{\cdot})p_{i'\cdot}^2}{p_{\cdot\cdot}(p_{i'\cdot}+q_{i'})^2} \tag{63}$$

and

$$\frac{q_i}{p_{i\cdot}+q_i} = \frac{q_{i'}}{p_{i'\cdot}+q_{i'}}. \tag{64}$$

From (64), $q_i/p_{i\cdot} = q_{i'}/p_{i'\cdot}$, and hence $q_i = \zeta p_{i\cdot}$ $(\zeta > 0)$ for any i. By substituting these relations for (63), $\kappa = \zeta$. Therefore, (A1) $\Rightarrow$ (A5). □

7.2 Proof of Theorem B

Proof of Theorem B. The equivalence (B2) $\Leftrightarrow$ (B3) is obvious from (8). The chi-square distance $\tilde{\tau}_c^2(t, c+1)$ between the t-th and the $(c+1)$-th column variables based on the profile matrix $\tilde{\boldsymbol{P}}$ is expressed as

$$\tilde{\tau}_c^2(t,c+1) = \sum_{i}^{r}\frac{p_{\cdot\cdot}+q_{\cdot}}{p_{i\cdot}+q_i}\left(\frac{p_{it}}{p_{\cdot t}}-\frac{q_i}{q_{\cdot}}\right)^2.$$

By the definition of the chi-square distance, we obtain (B4) $\Leftrightarrow$ (B1).

Next we prove (B4) $\Rightarrow$ (B3). For some $t \in \{1, \dots, c\}$, the (t, j)-th element of $\tilde{\boldsymbol{\Theta}}$ is expressed as

$$\theta_{tj} = \sum_{i=1}^{r} \frac{\tilde{p}_{it}\tilde{p}_{ij}}{\tilde{p}_{i\cdot}\tilde{p}_{\cdot t}}, \tag{65}$$

where $\tilde{p}_{ij}$ is the (i, j)-th element of $\tilde{\boldsymbol{P}}$. For $t \in \{1, \dots, c\}$, $\tilde{p}_{\cdot t} = p_{\cdot t}$. Also, $\tilde{p}_{i(c+1)} = q_i$ and $\tilde{p}_{\cdot(c+1)} = q_{\cdot}$. By (B4), for $j = 1, \dots, c+1$,

$$\theta_{tj} = \sum_{i=1}^{r} \frac{p_{it}}{p_{\cdot t}}\frac{\tilde{p}_{ij}}{\tilde{p}_{i\cdot}} = \sum_{i=1}^{r} \frac{q_i}{q_{\cdot}}\frac{\tilde{p}_{ij}}{\tilde{p}_{i\cdot}} = \sum_{i=1}^{r} \frac{\tilde{p}_{i(c+1)}}{\tilde{p}_{\cdot(c+1)}}\frac{\tilde{p}_{ij}}{\tilde{p}_{i\cdot}} = \theta_{(c+1)j}. \tag{66}$$

Let $\tilde{\lambda}_\alpha$ and $\tilde{\boldsymbol{\mu}}_\alpha = (\tilde{\mu}_{1\alpha}, \dots, \tilde{\mu}_{(c+1)\alpha})'$ be the eigenvalue and eigenvector of $\tilde{\boldsymbol{\Theta}}$, respectively. From the relation $\tilde{\boldsymbol{\Theta}}\tilde{\boldsymbol{\mu}}_\alpha = \tilde{\lambda}_\alpha\tilde{\boldsymbol{\mu}}_\alpha$, the t-th and the $(c+1)$-th elements are expressed as

$$\sum_{j=1}^{c+1} \theta_{tj}\tilde{\mu}_{j\alpha} = \tilde{\lambda}_\alpha\tilde{\mu}_{t\alpha}, \qquad \sum_{j=1}^{c+1} \theta_{(c+1)j}\tilde{\mu}_{j\alpha} = \tilde{\lambda}_\alpha\tilde{\mu}_{(c+1)\alpha}, \tag{67}$$

respectively. Therefore, from (66), we have $\tilde{\mu}_{t\alpha} = \tilde{\mu}_{(c+1)\alpha}$ for $\tilde{\lambda}_\alpha \neq 0$.

Next we prove (B3) $\Rightarrow$ (B4). By (B3) and (67),

$$\sum_{j=1}^{c+1} \theta_{tj}\tilde{\mu}_{j\alpha} = \sum_{j=1}^{c+1} \theta_{(c+1)j}\tilde{\mu}_{j\alpha}.$$

Using (65), the above equation is written as

$$\sum_{i=1}^{r}\sum_{j=1}^{c+1} \left\{ \frac{\tilde{p}_{it}}{\tilde{p}_{\cdot t}} - \frac{\tilde{p}_{i(c+1)}}{\tilde{p}_{\cdot(c+1)}} \right\} \frac{\tilde{p}_{ij}}{\tilde{p}_{i\cdot}}\tilde{\mu}_{j\alpha} = 0.$$

From (52), for α $(\alpha = 1, \dots, d)$, we have

$$\sum_{i=1}^{r} \left\{ \frac{\tilde{p}_{it}}{\tilde{p}_{\cdot t}} - \frac{\tilde{p}_{i(c+1)}}{\tilde{p}_{\cdot(c+1)}} \right\} \tilde{v}_{i\alpha} = 0 \qquad \text{or} \qquad \tilde{\boldsymbol{v}}'_\alpha \tilde{\boldsymbol{P}} \tilde{\boldsymbol{\Delta}}_c^{-1} \boldsymbol{e}_{t(c+1)} = 0, \tag{68}$$

where $\boldsymbol{e}_{t(c+1)}$ is the $(c+1) \times 1$ vector which the t-th element is 1 and the $(c+1)$-th element is -1, the others 0. To get (B4), it is sufficient to show $\tilde{\tau}_c^2(t, c+1) = 0$. The last expression in (68) can be written as

$$\tilde{\boldsymbol{v}}'_\alpha \tilde{\boldsymbol{\Delta}}_r \boldsymbol{g} = 0 \qquad \text{for } \alpha = 1, \dots, \tilde{d},$$

where $\boldsymbol{g} = \tilde{\boldsymbol{\Delta}}_r^{-1} \tilde{\boldsymbol{P}} \tilde{\boldsymbol{\Delta}}_c^{-1} \boldsymbol{e}_{t(c+1)}$. This implies that $\boldsymbol{g} = \boldsymbol{0}$ or $\kappa \boldsymbol{1}_r$ (κ is a constant). If $\boldsymbol{g} = \kappa \boldsymbol{1}_r$, we get $0 = \boldsymbol{1}' \tilde{\boldsymbol{\Delta}}_r \boldsymbol{g} = \kappa$, which gives a contradiction. Therefore, $\boldsymbol{g} = 0$, and hence $\boldsymbol{e}'_{t(c+1)} \tilde{\boldsymbol{\Delta}}_c^{-1} \tilde{\boldsymbol{P}}' \boldsymbol{g} = \tilde{\tau}_c^2(t, c+1) = 0$. $\square$

7.3 Proofs of Theorem A$'$ and B$'$

PROOF OF THEOREM A$'$. The implication (A2$'$) $\Rightarrow$ (A1$'$) is easily derived as a generalization of the proof of Theorem A. Therefore we prove (A1$'$) $\Rightarrow$ (A2$'$). For any i and i', the chi-square distance $\tilde{\tau}_r^2(i,i')$ between the i-th and the i'-th row variables is expressed as

$$\tilde{\tau}_r^2(i,i') = \frac{p_{\cdot\cdot}+q_{\cdot\cdot}}{p_{\cdot\cdot}}\left\{\frac{p_{i\cdot}^2\gamma_{ii}}{(p_{i\cdot}+q_{i\cdot})^2} - \frac{2p_{i\cdot}p_{i'\cdot}\gamma_{ii'}}{(p_{i\cdot}+q_{i\cdot})(p_{i'\cdot}+q_{i'\cdot})} + \frac{p_{i'\cdot}^2\gamma_{i'i'}}{(p_{i'\cdot}+q_{i'\cdot})^2}\right\}$$

$$+\sum_{j=1}^{c_1}\frac{p_{\cdot\cdot}+q_{\cdot\cdot}}{q_{\cdot j}}\left(\frac{q_{ij}}{p_{i\cdot}+q_{i\cdot}} - \frac{q_{i'j}}{p_{i'\cdot}+q_{i'\cdot}}\right)^2,$$

where $\gamma_{ii'}$ is the one given in the proof of Theorem A. Since $\tilde{\tau}_r^2(i,i') = \dfrac{1}{\kappa+1}\tau_r^2(i,i')$,

$$\frac{1}{\kappa+1} = \frac{(p_{\cdot\cdot}+q_{\cdot\cdot})p_{i\cdot}^2}{p_{\cdot\cdot}(p_{i\cdot}+q_{i\cdot})^2} = \frac{(p_{\cdot\cdot}+q_{\cdot\cdot})p_{i\cdot}p_{i'\cdot}}{p_{\cdot\cdot}(p_{i\cdot}+q_{i\cdot})(p_{i'\cdot}+q_{i'\cdot})} = \frac{(p_{\cdot\cdot}+q_{\cdot\cdot})p_{i'\cdot}^2}{p_{\cdot\cdot}(p_{i'\cdot}+q_{i'\cdot})^2}$$

and

$$\frac{q_{ij}}{p_{i\cdot}+q_{i\cdot}} = \frac{q_{i'j}}{p_{i'\cdot}+q_{i'\cdot}},$$

where $\kappa = \sum_{j=1}^{c_1}\kappa_j$. Let $q_{ij} = (p_{i\cdot}+q_{i\cdot})\zeta_j$ for any i. Then $q_{i\cdot} = (p_{i\cdot}+q_{i\cdot})\zeta$, where $\zeta = \sum_{j=1}^{c_1}\zeta_j$, *i.e.*, $q_{i\cdot} = \dfrac{\zeta}{1-\zeta}p_{i\cdot}$. By substituting the result for $\dfrac{1}{\kappa+1} = \dfrac{(p_{\cdot\cdot}+q_{\cdot\cdot})p_{i\cdot}^2}{p_{\cdot\cdot}(p_{i\cdot}+q_{i\cdot})^2}$, $\dfrac{1}{\kappa+1} = 1-\zeta$. Therefore $q_{i\cdot} = \kappa p_{i\cdot}$. That is,

$$\sum_{j=1}^{c_1}(q_{ij} - \kappa_j p_{i\cdot}) = 0.$$

Also, since q_{ij} can be expressed as $\dfrac{\zeta_j}{1-\zeta}p_{i\cdot}$, q_{ij} is proportional to $p_{i\cdot}$ with respect to j. Therefore, we have $q_{ij} = \kappa_j p_{i\cdot}$. □

PROOF OF THEOREM B$'$. In all cases, since the proofs are easily derived as a generalization of the proof of Theorem B, they are omitted. □

Acknowledgments

I wish to express my deepest gratitude to Prof. Y. Fujikoshi of Hiroshima University for his constant comments and useful advices. Also I would like to thank Prof. M. Ohtaki of Hiroshima University for introducing me various fields of applied statistics, and to be grateful to Dr. K. Naito of Shimane University, Dr. H. Fujisawa of Tokyo Institute of Technology and Dr. K. Satoh of Hiroshima University for their encouragements and advices. Further, I also thank the referee for some very useful suggestions.

References

[1] A. Agresti, Categorical Data Analysis, John Wiley & Sons, 1990.
[2] M. L. Eaton and D. Tyler, The asymptotic distribution of singular values with applications to canonical correlation and correspondence analysis. J. Multivariate Anal. **50** (1994), 238–264.
[3] Y. Fujikoshi, A test for additional information in canonical correlation analysis. Ann. Inst. Statist. Math. **34** (1982), 523–530.
[4] Y. Fujikoshi, Tests for redundancy of some variables in multivariate analysis, In: Statistical Data Analysis and Inference (Y. Dodge, ed.), North-Holland, 1989.
[5] F. A. Graybill, Introduction to Matrices with Applications in Statistics, Wadsworth, 1969.
[6] M. J. Greenacre, Theory and Applications of Correspondence Analysis, Academic Press, 1984.
[7] S. J. Haberman, The Analysis of Frequency Data, Univ. of Chicago Press, Chicago, 1974.
[8] M. O. Hill, Correspondence analysis: a neglected multivariate method. J. Roy. Statist. Soc. Ser. C **23** (1974), 340–354.
[9] J. C. Hsu, Multiple Comparisons: Theory and methods, Chapman & Hall, 1996.
[10] N. L. Johnson and S. Kotz, Distributions in Statistics: Continuous Multivariate Distributions, John Wiley & Sons, 1972.
[11] L. Lebart, A. Morineau and K. M. Warwick, Multivariate Descriptive Statistical Analysis, John Wiley & Sons, 1984.
[12] T. Nakayama, K. Naito and Y. Fujikoshi, Stability of correspondence analysis and its alternative using Hellinger distance for contingency table. Internat. J. Math. Statist. Sci. **7** (1998), 97–119.
[13] C. R. Rao, Tests with discriminant functions in multivariate analysis. Sankhyā **7** (1946), 407–414.
[14] C. R. Rao, Tests of significance in multivariate analysis. Biometrika **35** (1948), 58–79.
[15] C. R. Rao, Inference on discriminant function coefficients, In: Essays in probability and statistics (R. C. Bose and Others, ed.), Univ. of North Carolina Press, 1970.
[16] C. R. Rao, Linear Statistical Inference and Its Applications, John Wiley & Sons, 1973.
[17] C. R. Rao, The use of Hellinger distance in graphical displays of contingency table data, In: New Trends in Probability and Statistics Volume 3 (M. Tiit *et. al.* Ed.), VSP/TEV, 1995.
[18] C. R. Rao, An alternative to correspondence analysis using Hellinger distance. Proc. Int. Symp. On Contemporary Multivariate Analysis and its Applications, Hong Kong, A11–A29, 1997.

Department of Mathematics
Faculty of Science
Hiroshima University
Higashi-Hiroshima 739-8526, Japan

Current address: Radiation Risk Analysis Laboratory
Department of Health Physics
Japan Atomic Energy Research Institute
Tokai-mura Ibaraki-ken 319-1195, Japan
nakayama@riskest.tokai.jaeri.go.jp