VERSATILE TESTS FOR NON-LINEAR DATA IN $ 2 ¥times mathrm{k} $ TABLES

Jayasekara, Leslie
Department of Mathematics, University of Ruhuna

Nishiyama, Harutoshi
Ariake National College of Technology

Yanagawa, Takashi
Graduate School of Mathematics, Kyushu University

https://doi.org/10.5109/13489

出版情報：Bulletin of informatics and cybernetics. 31 (2), pp.179-190, 1999-12. Research Association of Statistical Sciences
バージョン：
権利関係：

VERSATILE TESTS FOR NON-LINEAR DATA IN 2 × k TABLES

By

Leslie JAYASEKARA *, Harutoshi NISHIYAMA †, and Takashi YANAGAWA ‡

Abstract

Properties of the Q_t-test which was proposed in Jayasekara and Yanagawa (1995) for detecting non-linear differences of two populations in an ordinal categorical table are further studied in this paper under different asymptotic framework. It is assumed in this paper that each marginal sum relative to the gland total tends to constant which is away from zero when the gland total tends to infinitive. The asymptotic distributions of the test statistics are obtained under null and contigious alternatives. It is shown that single test statistic, or combinatin of several test statistics have high powers for detecting various patterns of non-linear responses.

Key Words and Phrases: location-dispersion test; Wilcoxon test; Nair's dispersion test; Mantel's extended test; Gram-Schmidit orthonormalization; cumulative chi-squared test.

1. Introduction

Conventionally, the Wilcoxon test(1945), or equivalently Mantel's extended test(1963), which are often called the U-test, has been applied for testing diference of two populations in ordered categorical data in $2 \times k$ tables. The test has high powers for detecting linear, or log linear responses, but poorly behaves for detecting non-linear response which we are interested in in this paper. The cumulative chi-square test (Takeuchi and Hirotsu, 1982; Hirotsu, 1983; Nair, 1987) and Nair's test (1986) have been developed for non-linear responses. The former test is an omnibus test developed for a wider class of responses and the latter test was, in particular, designed to detect the dispersion alternatives. Jayasekara, Yanagawa and Tsujitani (1994) developed a location-dispersion test and Jayasekara and Yanagawa(1995) extended it for detecting location, dispersion and higher order differences of two populations. The test is called the Q_t-test and its usefulness was shown in their paper for such $2 \times k$ tables that one marginal sum dominates all the the others. The test statistics of the Q_t-test is defined as partial sum of test statistics that are systematically constructed by applying the Gram-Schmidit orthonormalization to the Wilcoxon score vector. We show in this paper further properties of the test statistics under the different asymptotic framework from that considered in Jayasekara and

* Department of Mathematics, University of Ruhuna, Matara, Sri Lanka

† Ariake National College of Technology, Omuta, Fukuoka 836-8585, Japan

‡ Graduate School of Mathematics, Kyushu University 33, Fukuoka 812-8581, Japan

Yanagawa(1995). That is, it is assumed in this paper that each marginal sum relative to the gland total tends to constant which is away from zero when the gland total tends to infinitive. The asymptotic distributions of the test statistics are obtained under null and contigious alternatives. It is suggested that single component of Q_t-test statistic, or combinatin of several components might have high powers for detecting various patterns of non-linear responses.

2. The Test Statistics

2.1. Statistics based on the Wilcoxon score

Consider $2 \times k$ table given in Table 1, and suppose that $\mathbf{Y}_1 = (Y_{11}, Y_{12}, \cdots, Y_{1k})'$ and $\mathbf{Y}_2 = (Y_{21}, Y_{22}, \cdots, Y_{2k})'$ are independently distributed multinomial random vectors with parameters n_1, $(p_{11}, \cdots, p_{1k})$ and n_2, $(p_{21}, \cdots, p_{2k})$, respectively. We consider the following null hypothesis:

H_0:$\mathbf{Y}_1$ and $\mathbf{Y}_2$ are identically distributed.

Table 1: $2 \times k$ contingency table.

	Ordered Categories						Total
Group 1	Y_{11}	Y_{12}	.	.	.	Y_{1k}	n_1
Group 2	Y_{21}	Y_{22}	.	.	.	Y_{2k}	n_2
Total	τ_1	τ_2	.	.	.	τ_k	N

To obtain test statistics for H_0 against non-linear alternatives, the following orthonormal scores based on the Wilcoxon score was introduced in Jayasekara and Yanagawa (1995). Let c_i be the Wilcoxon Score defined by

$$c_i = \sum_{j=1}^{i-1} \tau_j + (\tau_i - N)/2, \quad (i = 1, 2, \cdots, k)$$

so that $\sum_{i=1}^{k} \tau_i c_i = 0$. For k dimensional vectors $\mathbf{a} = (a_1, a_2, \cdots, a_k)'$ and $\mathbf{b} = (b_1, b_2, \cdots, b_k)'$ the inner product of $\mathbf{a}$ and $\mathbf{b}$ is defined by

$$(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{k} \tau_i a_i b_i$$

and the norm by $\|\mathbf{a}\|^2 = (\mathbf{a}, \mathbf{a})$.

Let c_j^i be the i-th power of c_j, $j = 1, 2, \cdots, k$ and put

$$\mathbf{c}_i = (c_1^i, c_2^i, \cdots, c_k^i)', \; i = 0, 1, 2, \cdots, k-1.$$

In particulra, $\mathbf{c}_0 = (1, 1, \cdots, 1)$. It is obvious that $\mathbf{c}_0, \mathbf{c}_1, \cdots, \mathbf{c}_{k-1}$ are linearly independent. Let $\mathbf{a}_0, \mathbf{a}_1, \cdots, \mathbf{a}_{k-1}$ be orthonormal score vectors which are obtained by applying Gram-Schmidit orthonormalization to these vectors; that is,

$$\mathbf{a}_0 = \mathbf{c}_0/\|\mathbf{c}_0\| \quad \text{and} \quad \mathbf{a}_r = \mathbf{d}_r/\|\mathbf{d}_r\|, \quad r = 1, 2, \cdots, k-1,$$

where

$$\mathbf{d}_r = \mathbf{c}_r - \sum_{l=0}^{r-1} (\mathbf{c}_r, \mathbf{a}_l)\mathbf{a}_l.$$

We have,

$$(\mathbf{a}_i, \mathbf{a}_j) = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases} \tag{2.1}$$

Representing the components of $\mathbf{a}_r$ by $\mathbf{a}_r = (a_{r1}, a_{r2}, \cdots, a_{rk})'$, Jayasekara and Yanagawa (1995) considered

$$U_r = \sum_{i=1}^{k} a_{ri} Y_{2i} / \sqrt{n_1 n_2 / N(N-1)}, \quad \text{for } r = 1, 2, \cdots, k-1.$$

2.2. Characteristic of the statistics

Now we consider the characteritic of U_r. We first remind that, under H_0, the conditional distribution of $\mathbf{Y}_2$ conditioned on $C = \{n_1, n_2, \tau_1, \cdots, \tau_k\}$ is multiple hypergeometric with

$$\begin{aligned} E[Y_{2j}|C] &= n_2 \tau_j / N \qquad (2.2) \\ Cov[Y_{2j}, Y_{2j'}|C] &= \frac{n_1 n_2}{N^2(N-1)} \tau_j (\delta_{jj'} N - \tau_{j'}), \quad \text{for } j, j' = 1, 2, \cdots, k, \end{aligned}$$

where $\delta_{jj'} = 1$ if $j = j'$ and 0 otherwise.

Put $\mathbf{Y}_2 = (Y_{21}, Y_{22}, \ldots, Y_{2k})'$ and

$$\mathbf{U} = (U_1, U_2, \ldots, U_t)' = \frac{\mathbf{A}\mathbf{Y}_2}{\sqrt{n_1 n_2 / N(N-1)}}, \tag{2.3}$$

where $\mathbf{A} = (a_{rj})$ is the $t \times k$ matrix.

We have the following theorems.

THEOREM 2.1. *Under H_0, $U_r, r = 1, 2, \cdots, t$, are uncorrelated with zero mean and unit variance.*

PROOF. From (2.1) we have

$$\mathbf{A}(\tau_1, \cdots, \tau_k)' = \mathbf{0} \tag{2.4}$$

$$\mathbf{A} \begin{pmatrix} \tau_1 & & & \\ & \cdot & & \mathbf{0} \\ & \mathbf{0} & \cdot & \\ & & & \tau_k \end{pmatrix} \mathbf{A}' = \mathbf{I}_t \tag{2.5}$$

Using (2.2) and (2.3) we have,

$$
\begin{aligned}
E[\mathbf{U}] &= \frac{\mathbf{A}}{\sqrt{n_1 n_2/N(N-1)}} E[\mathbf{Y}_2] \\
&= \frac{n_2/N}{\sqrt{n_1 n_2/N(N-1)}} A(\tau_1, \cdots, \tau_k)' \\
&= \mathbf{0}.
\end{aligned}
$$

From (2.3) and (2.4) covariance matrix of $\mathbf{U}$ is

$$
\begin{aligned}
Cov(\mathbf{U}) &= \frac{\mathbf{A} Cov(\mathbf{Y}_2)\mathbf{A}'}{n_1 n_2/N(N-1)} \\
&= \mathbf{A} \begin{pmatrix} \tau_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \tau_k \end{pmatrix} \mathbf{A}' - \frac{1}{N}\mathbf{A}(\tau_1, \cdots, \tau_k)'(\tau_1, \cdots, \tau_k)\mathbf{A}' \\
&= \mathbf{I}_t.
\end{aligned}
$$

□

The proofs of the following theorems will be given in Section 3. Note that the limiting condition considered in Jayasekara and Yanagawa (1995) was $n_i/N \to 0$ for $i = 2, 3, \ldots, k$ when $N \to \infty$.

THEOREM 2.2. *Suppose that* $N^{1/2}(n_i/N - \gamma_i) \to 0$, $(0 < \gamma_i < 1;\ i = 1, 2)$ *and* $N^{1/2}(\tau_j/N - \rho_j) \to 0$ $(0 < \rho_j < 1;\ j = 1, 2, \cdots, k)$ *as* $N \to \infty$. *Then under* H_0 $\mathbf{U}$ *follows asymptotically t dimensional normal distribution with mean vector* $\mathbf{0}$ *and identity covariance matrix as* $N \to \infty$.

Putting $\psi_j = p_{11}p_{2j}/p_{21}p_{1j}$ $(j = 1, 2, \cdots, k)$ so that $\psi_1 = 1$, the asymptotic distribution of $\mathbf{U}$ under alternative hypothesis

$$H_1 : \psi_j = 1 + A_j/N^{1/2},\ j = 2, 3, \cdots, k$$

is given in the following theorem, where A_j is a constant.

THEOREM 2.3. *Assume the same condition as Theorem 2.2, then under* H_1 $\mathbf{U}$ *follows asymptotically t dimensional normal distribution with mean vector* $\boldsymbol{\delta}$ *and identity covariance matrix as* $N \to \infty$, *where the r-th component of* $\boldsymbol{\delta}$ *is given by*

$$\delta_r = \left(\frac{(N-1)n_1 n_2}{N}\right)^{1/2}(\mathbf{a}_r, \boldsymbol{\psi} - \mathbf{1}),$$

where $\boldsymbol{\psi} = (\psi_1, \psi_2, \cdots, \psi_k)'$ *and* $\mathbf{1} = (1, \cdots, 1)'$.

Since $\log \psi_j \sim \psi_j - 1$ when ψ_j is close to 1 and inner product $(\mathbf{a}_r, \log \boldsymbol{\psi})$ is maximized when $\log \boldsymbol{\psi} = \beta \mathbf{a}_r$ we have the following corollary.

COROLLARY 2.4. *The asymptotic power of the test based on statistic U_r is maximized when* $\log\psi = \beta\mathbf{a}_r$, *where β is a scalar constant.*

We also have the following theorem.

THEOREM 2.5.

$$\sum_{r=1}^{k} U_r^2 = \frac{N-1}{N}\sum_{i=1}^{2}\sum_{j=1}^{k}\frac{(O-E)^2}{E},$$

where $O = Y_{ij}$ and $E = n_i\tau_j/N$; that is $\sum_{r=1}^{k} U_r^2$ is equivalent to the Pearson chi-squared test statistic (1990) except for constant $(N-1)/N$.

PROOF. Note first that

$$\sum_{i=1}^{2}\sum_{j=1}^{k}\frac{(O-E)^2}{E} = \frac{N^2}{n_1 n_2}\sum\frac{1}{\tau_j}(Y_{2j} - \frac{\tau_j n_2}{N})^2. \tag{2.6}$$

Now put $\boldsymbol{\tau}=(\tau_1,\ldots,\tau_k)'$, then since $\mathbf{AY}_2 = \mathbf{A}(\mathbf{Y}_2 - \boldsymbol{\tau}n_2/N)$ we have

$$\begin{aligned}(\mathbf{AY}_2)'\mathbf{AY}_2 &= (\mathbf{A}(\mathbf{Y}_2 - \boldsymbol{\tau}n_2/N))'\mathbf{A}(\mathbf{Y}_2 - \boldsymbol{\tau}n_2/N)\\ &= (\mathbf{Y}_2 - \boldsymbol{\tau}n_2/N)'\begin{pmatrix}1/\tau_1 & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & 1/\tau_k\end{pmatrix}(\mathbf{Y}_2 - \boldsymbol{\tau}n_2/N).\end{aligned}$$

□

2.3. Detecting for non-linear response

Figure 1(1)shows the patterns of $\log\psi$ that provide the maximum asymptotic powers to statistics U_1, U_2, U_3, U_4, and U_5, obtained from the corollary, when $k = 5$ and $\tau_1 = \tau_2 = ... = \tau_5 = 10$. Reflecting prior experience on response pattern one may select one of U_r's for powerful detection of difference. However, the patterns depend on the values of τ's; for example, see Figure 1(2) which shows substantially different patterns from Figure 1(1). If this is the case omnibus test such as

$$Q_t = \sum_{r=1}^{t} U_r^2,$$

(Jayasekara and Yanagawa, 1995), or multiple comparison test such as

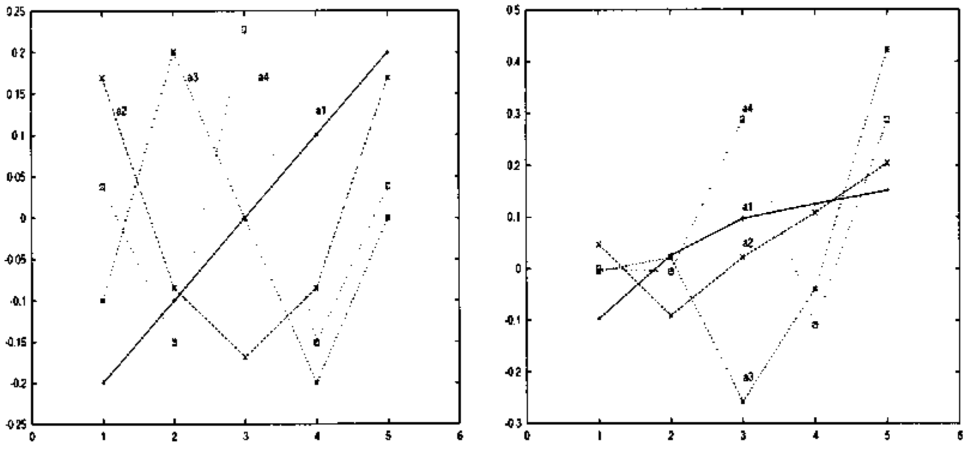
$$M_t = max\{|U_1|, |U_2|, \cdots, |U_t|\}$$

for each $t \in \{1, 2, \cdots, k-1\}$ might be useful.

Theorem 2.5 shows that if t is close to k Q_t-test behaves like the Pearson chi-squared test. Furthermore, it is not clear against which pattern of non-linear responses the Q_t-test is powerful. On the other hand M_t-test is a multiple comparison test, comparing t hypotheses

Figure 1: Patterns of Score Vectors

(1) When $\tau_1 = \cdots = \tau_5 = 10$ (2) For data given in Table 3.



$$H_0\colon \log\psi=(1,1,\ldots,1)',\mathrm{H}_1\colon \log\psi = \beta\mathbf{a}_1 \quad (\beta \neq 0),\ \ldots,\ H_t\colon \log\psi = \beta\mathbf{a}_t \quad (\beta \neq 0).$$

Thus if rejected by the M_t-test we may know which patterns of response is responsible for the rejection. By Theorem 2.2 it is straightfoward to obtain approximate critical points of the M_t-test. Those selected values are listed in Table 2.

Table 2: Selected Critical Points for the M_t-test (two-sided)

	α		
t	0.10	0.05	0.01
1	1.644	1.960	2.576
2	1.948	2.235	2.807
3	2.113	2.387	2.929
4	2.226	2.489	3.011
5	2.312	2.569	3.039

2.4. An application

Table 3 lists the efficacy of certain drug obtained at Phase III randomized clinical trial from 72 patients (Study 1) and that of the same drug obtained at a post market study from 73 patients (Study 2). CR, PR, MR, NR, and PD stand for completely recovered, partially recovered, moderately recovered, no recovered, and progressive of disease, respectively. The score vectors and values of U's which are computed from the table are as follows:

$\mathbf{a}_1$=(-0.097, 0.026, 0.097, 0.125, 0.151), $\mathbf{a}_2$=(0.047, -0.091, 0.022, 0.108, 0.204), $\mathbf{a}_3$=(-0.005, 0.022, -0.258, -0.039, 0.423), $\mathbf{a}_4$=(0.001, -0.006, 0.289, -0.108, 0.288),

U_1=-1.52, U_2=-1.97, U_3=2.11, and U_4=0.27.

Figure 1(2) shows the patterns of the score vectors. It is shown from Table 2 that when $t = 1$, the M_t-test results in no significance at 10% level; when t=2 the test detects such pattern of the log odds ratio as illustrated by the broken line in Figure 1(2) at 10% level; and when t=3 the M_t-test almost detects such pattern of the log odds ratio as illustrated by the dotted line in Figure 1(2) at 10% significant level. On the other hand the values of the Q_t statistics and their p-values are

Q_1=2.32 (p=0.128), Q_2=6.21 (p=0.045), Q_3=10.67 (p=0.014)

and Q_t-test detects non-linearlity difference better than M_t-test in the present example, but keeping silent about what pattern of non-linear response it detected.

Table 3: Efficacy of a Drug between Phase III and Post Market Studies

	CR	PR	MR	NR	PD	Total
Group 1	1	16	5	25	26	72
Group 2	2	5	1	36	28	73
Total	3	21	6	61	54	145

3. Proofs of Theorems 2.2 and 2.3

We use the normal approximation of a multiple hypergeometric distribution discussed in Plackett (1981), and briefly skech it in the next subsection.

3.1. The normal approximation of multiple hypergeometric distribution

When $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are independently distributed multinomial random vectors with the parameters n_1,$(p_{11}, \cdots, p_{1k})$ and n_2,$(p_{21}, \cdots, p_{2k})$, respectively, we have,

$$\Pr[(Y_{21}, \cdots, Y_{2k}) = (y_{21}, \cdots, y_{2k})|C] = \frac{\psi_1^{y_{21}} \cdots \psi_k^{y_{2k}} / \prod_{i=1}^2 \prod_{j=1}^k y_{ij}!}{\sum_{r_{21}+\cdots+r_{2k}=n_2} \psi_1^{r_{21}} \cdots \psi_k^{r_{2k}} / \prod_{i=1}^2 \prod_{j=1}^k r_{ij}!}, \tag{3.1}$$

where ψ_j is the odds ratio of j-th category with respect to category 1, i.e.

$$\psi_j = p_{11}p_{2j}/p_{21}p_{1j}, \quad (j = 1, 2, \cdots, k).$$

Note that $\psi_1 \equiv 1$. We use the following assumption.

Assumption (A1) As $N \to \infty$, $n_i/N \to \gamma_i, 0 < \gamma_i < 1$, for $i = 1, 2$, and $\tau_j/N \to \rho_j, 0 < \rho_j < 1$, for $j = 1, 2, \cdots, k$.

LEMMA 3.1. (Sinkhorn(1967)) *If $\{m_{ij}\}$ satisfy $\sum_{j=1}^{k} m_{ij} = n_i$, for $i = 1, 2$; $\sum_{i=1}^{2} m_{ij} = \tau_j$; and $(m_{11}m_{2j})/(m_{21}m_{1j}) = \psi_j$ for $j = 1, 2, \cdots, k$, then $\{m_{ij}\}$ are uniquely determined by the following algorithm;*

$$
\begin{aligned}
m_{1j}^{(1)} &= \frac{n_1}{k}, j = 1, 2, \cdots, k, \\
m_{21}^{(1)} &= \frac{n_2}{k[1 + \sum_{j=2}^{k}(\psi_j - 1)/k]} \\
m_{2j'}^{(1)} &= \frac{n_2 \psi_{j'}}{k[1 + \sum_{j=2}^{k}(\psi_j - 1)/k]}, j' = 2, \cdots, k, \\
m_{ij}^{(2)} &= \frac{m_{ij}^{(1)} \tau_j}{m_{.j}^{(1)}}, \\
m_{ij}^{(3)} &= \frac{m_{ij}^{(2)} n_i}{m_{i.}^{(2)}},
\end{aligned}
$$

$$
\begin{aligned}
m_{ij}^{(2h)} &= \frac{m_{ij}^{(2h-1)} \tau_j}{m_{.j}^{(2h-1)}}, \\
m_{ij}^{(2h+1)} &= \frac{m_{ij}^{(2h)} n_i}{m_{i.}^{(2h)}}, h = 1, 2, \cdots .
\end{aligned}
$$

THEOREM 3.2. (Plackett(1981)) *Suppose (A1), then the conditional distribution of $\mathbf{Y} = (Y_{22}, \cdots, Y_{2k})'$ conditioned on C converges in distribution to $N_{k-1}(\mathbf{m}_2, \mathbf{\Sigma})$ as $N \to \infty$, where $\mathbf{m}_2 = (m_{22}, \cdots, m_{2k})'$ and $\mathbf{\Sigma}^{-1} = (\sigma_{ij})$, and $\sigma_{ij} = m_{11}^{-1} + m_{21}^{-1} + (m_{1j}^{-1} + m_{2j}^{-1})\delta_{ij}$, $i, j = 2, \cdots, k$, where m_{ij}'s are quantities determined in Lemma 3.1.*

3.2. Evaluation of the scores

We need several lemmas to evaluate the convergent order of scores. We define $N^{-r}c_{ri} = O(1)$ if and only if $N^{-r}c_{ri}$ tends to a constant as $N \to \infty$.

LEMMA 3.3. *If (A1) is satisfied, then*

$$N^{-r}c_{ri} = O(1), (i = 1, 2, \cdots, k),$$

where $c_{ri} = {c_i}^r$, is the r-th power of the i-th Wilcoxon score.

PROOF. The lemma is proved if we may show $N^{-1}c_i = O(1)$. But since $c_i = \sum_{j=1}^{i-1} \tau_j + (\tau_i - N)/2$ we get $N^{-1}c_i = O(1)$ easily from (A1). □

LEMMA 3.4. *If (A1) is satisfied, then*

$$N^{-r}(\mathbf{c}_r, \mathbf{a}_0)a_{0i} = O(1), \quad r = 1, 2, \cdots, k-1, \quad i = 1, 2, \cdots, k.$$

PROOF. From the definition of $\mathbf{a}_0$ $a_{0i} = 1/N^{1/2}$ for all i. So by Lemma 3.3 we obtain $N^{-(r+1/2)}(\mathbf{c}_r, \mathbf{a}_0) = O(1)$. Hence, the desired result follows. □

LEMMA 3.5. *If (A1) is satisfied, then*

$$N^{-r}d_{ri} = O(1), \quad r = 1, 2, \cdots, k-1, \quad i = 1, 2, \cdots, k.$$

PROOF. To prove this result we use induction on r.
In case of $r = 1$,

$$d_{1i} = c_{1i} - (c_1, a_0)a_{0i}, \quad \text{for } i = 1, 2, \cdots, k.$$

Applying Lemma 3.3 and 3.4, it follows that

$$N^{-1}d_{1i} = O(1), \text{for } i = 2, 3, \cdots, k.$$

Suppose that the result is true for $r = 1, 2, \cdots, m-1$. We have

$$\begin{aligned}\mathbf{d}_m &= \mathbf{c}_m - \sum_{l=0}^{m-1}(\mathbf{c}_m, \mathbf{a}_l)\mathbf{a}_l, \\ &= \mathbf{c}_m - (\mathbf{c}_m, \mathbf{a}_0)\mathbf{a}_0 - \sum_{l=1}^{m-1}(\mathbf{c}_m, \mathbf{d}_l)\frac{\mathbf{d}_l}{\|\mathbf{d}_l\|^2},\end{aligned}$$

it follows that $N^{-m}c_{mi} = O(1)$ from Lemma 3.3, and $N^{-m}((\mathbf{c}_m, \mathbf{a}_0)\mathbf{a}_0 = O(1)$ from Lemma 3.4. Furthermore, since by the assumption of induction and Lemma 3.4 N^{-m-l} $(\mathbf{c}_m, \mathbf{d}_l) = O(1)$ for $l = 1, 2, \cdots, m-1$, and also from the assumption of induction we have $N^l\mathbf{d}_l/\|\mathbf{d}_l\|^2 = O(1)$ for $l = 1, 2, \cdots, m-1$. Therefore

$$N^{-m}(\mathbf{c}_m, \mathbf{d}_l)\frac{\mathbf{d}_l}{\|\mathbf{d}_l\|^2} = O(1).$$

So the result is true for $r = m$. By the induction the proof is completed. □

Using these lemmas we may easily show

LEMMA 3.6. $N^{1/2}a_{ri} = O(1), \quad for \quad i = 1, 2, \cdots, k, \quad r = 1, 2, \cdots, k-1.$

3.3. Proof of Theorem 2.2

LEMMA 3.7. *Assume (A1), then under H_0 the conditional distribution of $N^{-1/2}(\mathbf{Y} - N\gamma_2\boldsymbol{\rho})$ conditioned on C converges in distribution to $N_{k-1}(\mathbf{0}, \boldsymbol{\Sigma}_0)$ as $N \to \infty$, where $\boldsymbol{\rho} = (\rho_2, \cdots, \rho_k)'$ and ${\boldsymbol{\Sigma}_0}^{-1} = (\sigma_{ij0})$ with $\sigma_{ij0} = [\rho_1^{-1} + \delta_{ij}\rho_j^{-1}]/\gamma_1\gamma_2$, $i, j = 2, \cdots, k$.*

PROOF. Under H_0, $m_{ij} = N\gamma_i\rho_j$, for $i = 1, 2$, and $j = 1, \cdots, k$. Thus from Theorem 3.2 we have the desired result. □

Now assuming stronger assumption than (A1)
Assumption (A2) $N^{1/2}(n_i/N - \gamma_i) \to 0$, $(0 < \gamma_i < 1; i = 1, 2)$ and $N^{1/2}(\tau_j/N - \rho_j) \to 0$ $(0 < \rho_j < 1; j = 1, 2, \cdots, k)$ as $N \to \infty$
we prove Theorem 2.2. Put $\mathbf{B} = (a_{rj} - a_{r1})$, $r = 1, 2, \cdots, t$, and $j = 2, 3, \cdots, k$. Then from (2.3) and (A2) we may represent

$$\begin{aligned} \mathbf{U} &= \frac{\mathbf{A}\mathbf{Y}_2}{\sqrt{n_1 n_2/N(N-1)}} \\ &= \mathbf{B}(\mathbf{Y} - N\gamma_2\boldsymbol{\rho})/\sqrt{n_1 n_2/N(N-1)} + o(1) \\ &= (N^{1/2}\mathbf{B})N^{-1/2}(\mathbf{Y} - N\gamma_2\boldsymbol{\rho})/\sqrt{n_1 n_2/N(N-1)} + o(1) \end{aligned}$$

Thus from Lemma 3.6 and Lemma 3.7 **U** converges in distribution to a multivariate normal distribution. Finally from Theorem 2.1 it follows that U has mean vector **0** and identity covariance matrix. This completes the proof of Theorem 2.2.

Remark: As a by-product of the above proof it follows that

$$\mathbf{B}\Sigma_0\mathbf{B}'/\{n_1 n_2/N^2(N-1)\} = I_t + o(1).$$

3.4. Proof of Theorem 2.3

Recall that the alternative hypothesis we consider is

$$H_1 : \psi_j = 1 + A_j/N^{1/2}, j = 2, \cdots, k.$$

LEMMA 3.8. *Under H_1, the quantities in Theorem 3.2 are given by*

$$\begin{aligned} m_{1j}^{(1)} &= \frac{n_1}{k} \\ m_{21}^{(1)} &= \frac{n_2}{k}[1 - \sum_{j=2}^{k}\frac{(\psi_j - 1)}{k} + o(N^{-1/2})] \\ m_{2j'}^{(1)} &= \frac{n_2}{k}[\psi_{j'} - \sum_{j=2}^{k}\frac{(\psi_j - 1)}{k} + o(N^{-1/2})], \quad j' = 2, 3, \cdots, k, \\ m_{i1}^{(2)} &= N\gamma_i\rho_1 + (-1)^{i+1}N\gamma_1\gamma_2\rho_1\sum_{j=2}^{k}\frac{(\psi_j - 1)}{k} + o(N^{1/2}), \\ m_{ij'}^{(2)} &= N\gamma_i\rho_{j'} + (-1)^{i}N\gamma_1\gamma_2\rho_{j'}[\psi_{j'} - 1 - \sum_{j=2}^{k}\frac{\psi_j - 1}{k}] + o(N^{1/2}), \\ m_{ij}^{(l)} &= N\gamma_i\rho_j + N^{1/2}\eta_{ij} + o(N^{1/2}), \end{aligned}$$

for $l = 3, 4, \cdots$, *where*

$$\eta_{i1} = (-1)^{i+1} N^{1/2} \gamma_1 \gamma_2 \rho_1 \sum_{j=2}^{k} (\psi_j - 1)\rho_j,$$

$$\eta_{ij'} = (-1)^{i} N^{1/2} \gamma_1 \gamma_2 \rho_{j'} [\psi_{j'} - 1 - \sum_{j=2}^{k} (\psi_j - 1)\rho_j], \quad (j' = 2, 3, \cdots, k).$$

PROOF. Substituting $\psi_j = 1 + A_j/N^{1/2}$ to the algorithm which is given in Lemma 3.1 and using the Taylor expansion we easily have the expressions for $m_{ij}^{(1)}, m_{21}^{(1)}, m_{2j'}^{(1)}$, $m_{i1}^{(2)}$ and $m_{ij}^{(2)}$. Similarly substituting ψ_j and using mathematical induction on l, the final expansions can be obtained. □

From Lemma 3.8 we may represent m_{ij} under H_1 by

$$m_{ij} = N\gamma_i \rho_j + N^{1/2}\eta_{ij} + o(N^{1/2})$$

for $i = 1, 2$ and $j = 1, 2, \cdots, k$. It follows that the conditional distribution of $N^{-1/2}(\mathbf{Y} - N\gamma_2 \boldsymbol{\rho})$ conditioned on C coverges in distribution to $N_{k-1}(\eta_2, \boldsymbol{\Sigma}_0)$ as $N \to \infty$ under H_1, where $\eta_2 = (\eta_{22}, \cdots, \eta_{2k})'$. Now since

$$\mathbf{U} = (N^{1/2}\mathbf{B}) N^{-1/2} (\mathbf{Y} - N\gamma_2 \boldsymbol{\rho}) / \sqrt{n_1 n_2 / N(N-1)} + o(1),$$

the conditional distribution of $\mathbf{U}$ conditioned on C converged in distribution to $N_t(\mathbf{B}\boldsymbol{\eta}_2 \sqrt{n_1 n_2 / N^2(N-1)}, \mathbf{B}\boldsymbol{\Sigma}_0 \mathbf{B}' / \{n_1 n_2 / N^2(N-1)\})$. From the remark at the end of the previous subsection we have

$$\lim_{N\to\infty} \mathbf{B}\boldsymbol{\Sigma}_0 \mathbf{B}' / \{n_1 n_2 / N^2 (N-1)\} \to \mathbf{I}_t.$$

Furthermore, since

$$\sum_{j=1}^{k} a_{rj}\rho_j = 0, \quad \textit{and} \quad \sum_{j=1}^{k} \rho_j = 1$$

the r-th element of $\mathbf{B}\boldsymbol{\eta}_2$ is simply represented by

$$\delta_r = \sum_{j=1}^{k} (a_{rj} - a_{r1})\eta_{2j} = N^{1/2}\gamma_1\gamma_2 \sum_{j=1}^{k} a_{rj}\rho_j(\psi_j - 1).$$

This completes the proof of Theorem 2.3.

References

Hirotsu, C.(1983). Defining the pattern of association in two-way contingency tables, *Biometrika*, **70**, 579-589.

Jayasekara, L., Yanagawa, T. and Tsujitani, M.(1994). A location-dispersion test for $2 \times k$ tables, *Bulletin of Informatics and Cybernetics*, **26**,125-139.

Jayasekara, L. and Yanagawa, T. (1995). Testing ordered categorical data in $2 \times k$ tables by the statistics with orthonormal score vectors, *Bulletin of Informatics and Cybernetics*, **27**, 129-145.

Mantel, N. (1963). Chi-Square tests with one degree of freedom: extensions of the Mantel-Haenszel procedure, *Journal of American Statistical Association*, **58**, 690-700.

Nair, V.N.(1986). On testing in industrial experiments with ordered categorical data, *Technometrics*, **28**, 283-311.

Nair, V.N. (1987). Chi-squared-type tests for ordered alternatives in contingency tables, *Journal of the American Statistical Association*, **82**, 283-291.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable, in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, *Philos. Mag.*, **50**, 157-175.

Plackett, R. L. (1981). *The Analysis of Categorical Data*, London: Griffin.

Sinkhorn, R. (1967). Diagonal equivalence to matrices with prescribed row and column sums, *Amer. Math. Month.*, **74**, 402-5.

Takeuchi, K and Hirotsu, C.(1982). The cumulative chi-squares method against ordered alternatives in two-way contingency tables, *Reports of Statistical Application Research, Japanese Union of Scientists and Engineers*, **29**, 1-13.

Wilcoxon, Frank (1945). Individual comparisons by ranking methods, *Biometrics* **1**, 80-83.

Received October 3,1998
Revised November 19,1999