VERSATILE TESTS FOR NON-LINEAR DATA IN $ 2 \times \text{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

http://hdl.handle.net/2324/13489
VERSATILE TESTS FOR NON-LINEAR DATA
IN $2 \times k$ 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 dis-
tributions of the test statistics are obtained under null and contiguous
alternatives. It is shown that single test statistic, or combination of sev-
eral 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-Schmidt 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 difference 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-Schmidt 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 contiguous alternatives. It is suggested that single component of $Q_t$-test statistic, or combination 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 $Y_1 = (Y_{11}, Y_{12}, \ldots, Y_{1k})'$ and $Y_2 = (Y_{21}, Y_{22}, \ldots, Y_{2k})'$ are independently distributed multinomial random vectors with parameters $n_1, (p_{11}, \ldots, p_{1k})$ and $n_2, (p_{21}, \ldots, p_{2k})$, respectively. We consider the following null hypothesis:

$$H_0: Y_1 \text{ and } Y_2 \text{ are identically distributed.}$$

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

<table>
<thead>
<tr>
<th>Ordered Categories</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Y_{11}$</td>
<td>$Y_{12}$</td>
<td>$Y_{1k}$</td>
</tr>
<tr>
<td></td>
<td>$Y_{21}$</td>
<td>$Y_{22}$</td>
<td>$Y_{2k}$</td>
</tr>
<tr>
<td>Total</td>
<td>$\tau_1$</td>
<td>$\tau_2$</td>
<td>$\tau_k$</td>
</tr>
</tbody>
</table>

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

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

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

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

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

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

$$c_i = (c_1^i, c_2^i, \ldots, c_k^i)', \quad i = 0, 1, 2, \ldots, k - 1.$$  

In particular, $c_0 = (1, 1, \ldots, 1)$. It is obvious that $c_0, c_1, \ldots, c_{k-1}$ are linearly independent. Let $a_0, a_1, \ldots, a_{k-1}$ be orthonormal score vectors which are obtained by applying Gram-Schmidt orthonormalization to these vectors; that is,
Versatile tests for non-linear data

\[ a_0 = c_0 / \| c_0 \| \quad \text{and} \quad a_r = d_r / \| d_r \|, \quad r = 1, 2, \ldots, k - 1, \]

where

\[ d_r = c_r - \sum_{i=0}^{r-1} (c_r, a_i) a_i. \]

We have,

\[ (a_i, a_j) = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases} \quad (2.1) \]

Representing the components of \( a_r \) by \( a_r = (a_{r1}, a_{r2}, \ldots, a_{rk})' \), Jayasekara and Yana-
gawa (1995) considered

\[ U_r = \sum_{i=1}^{k} a_{r_i} Y_{2i} / \sqrt{n_1 n_2 / N(N-1)}, \quad \text{for } r = 1, 2, \ldots, k - 1. \]

2.2. Characteristic of the statistics

Now we consider the characteristic of \( U_r \). We first remind that, under \( H_0 \), the
conditional distribution of \( Y_2 \) conditioned on \( C = \{n_1, n_2, \tau_1, \ldots, \tau_k\} \) is multiple hyper-
geometric with

\[ E[Y_{2j} | C] = n_2 \tau_j / N \]

\[ \text{Cov}[Y_{2j}, Y_{2j'} | C] = \frac{n_1 n_2}{N^2(N-1)} \tau_j (\delta_{jj'} N - \tau_{j'}), \quad \text{for } j, j' = 1, 2, \ldots, k, \]

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

Put \( Y_2 = (Y_{21}, Y_{22}, \ldots, Y_{2k})' \) and

\[ U = (U_1, U_2, \ldots, U_t)' = \frac{AX_2}{\sqrt{n_1 n_2 / N(N-1)}}, \quad (2.3) \]

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

We have the following theorems.

Theorem 2.1. Under \( H_0 \), \( U_r, r = 1, 2, \ldots, t, \) are uncorrelated with zero mean and
unit variance.

Proof. From (2.1) we have

\[ A(\tau_1, \ldots, \tau_k)' = 0 \quad (2.4) \]

\[ A \begin{pmatrix} \tau_1 \\ 0 \\ \tau_k \end{pmatrix} A' = I_t \quad (2.5) \]
Using (2.2) and (2.3), we have:

\[
E[U] = \frac{A}{\sqrt{n_1n_2/N(N-1)}}E[Y_2] = \frac{n_2/N}{\sqrt{n_1n_2/N(N-1)}}A(\tau_1, \ldots, \tau_k)'
= 0.
\]

From (2.3) and (2.4), the covariance matrix of \( U \) is

\[
\text{Cov}(U) = \frac{ACov(Y_2)A'}{n_1n_2/N(N-1)} = A \begin{pmatrix} \tau_1 & \cdots & 0 \\ 0 & \cdots & \tau_k \end{pmatrix} A' - \frac{1}{N}A(\tau_1, \ldots, \tau_k)'(\tau_1, \ldots, \tau_k)A' = I_k.
\]

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 \( N^{1/2}(n_1/N - \gamma_1) \to 0 \), \( 0 < \gamma_1 < 1; i = 1, 2 \) and \( N^{1/2}(\tau_j/N - \rho_j) \to 0 \), \( 0 < \rho_j < 1; j = 1, 2, \ldots, k \) as \( N \to \infty \). Then under \( H_0 \), \( U \) follows asymptotically \( t \) dimensional normal distribution with mean vector \( 0 \) and identity covariance matrix as \( N \to \infty \).

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

\[
H_1: \psi_j = 1 + A_j/N^{1/2}, \ j = 2, 3, \ldots, k
\]

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 \( H_1 \) \( U \) follows asymptotically \( t \) dimensional normal distribution with mean vector \( \delta \) and identity covariance matrix as \( N \to \infty \), where the \( r \)-th component of \( \delta \) is given by

\[
\delta_r = \left(\frac{(N-1)n_1n_2}{N}\right)^{1/2}(a_r, \psi - 1),
\]

where \( \psi = (\psi_1, \psi_2, \ldots, \psi_k)' \) and \( 1 = (1, \ldots, 1)' \).

Since \( \log \psi_j \sim \tilde{\psi}_j - 1 \) when \( \psi_j \) is close to 1 and inner product \( (a_r, \log \psi) \) is maximized when \( \log \psi = \beta a_r \), we have the following corollary.
COROLLARY 2.4. The asymptotic power of the test based on statistic $U_r$ is maximized when $\log \psi = 3a_r$, where $3$ 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 $E = n_t \tau_j / N$; 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

$$\sum_{i=1}^{2} \sum_{j=1}^{k} \frac{(O - E)^2}{E} = \frac{N^2}{n_1 n_2} \sum_{j=1}^{\tau} (Y_{2j} - \frac{\tau_{n_2}}{N})^2.$$  \hspace{1cm} (2.6)

Now put $\tau = (\tau_1, \ldots, \tau_k)'$, then since $A Y_2 = A (Y_2 - \tau N_2/N)$ we have

$$(A Y_2)' A Y_2 = (A (Y_2 - \tau N_2/N))' A (Y_2 - \tau N_2/N)$$

$$= (Y_2 - \tau N_2/N)' \begin{pmatrix} 1/\tau_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1/\tau_k \end{pmatrix} (Y_2 - \tau N_2/N).$$

$\square$

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 $\tau_1 = \tau_2 = \ldots = \tau_5 = 10$. 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 $\tau$'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

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

for each $t \in \{1, 2, \ldots, k - 1\}$ 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.

$H_0: \log \psi = (1, 1, \ldots, 1)$, $H_1: \log \psi = \beta a_1 \ (\beta \neq 0)$, $\ldots$, $H_t: \log \psi = \beta a_t \ (\beta \neq 0)$.

Thus if rejected by the $M_5$-test we may know which patterns of response is responsible for the rejection. By Theorem 2.2 it is straightforward to obtain approximate critical points of the $M_5$-test. Those selected values are listed in Table 2.

Table 2: Selected Critical Points for the $M_5$-test (two-sided)

<table>
<thead>
<tr>
<th>$t$</th>
<th>0.10</th>
<th>0.05</th>
<th>0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.644</td>
<td>1.960</td>
<td>2.576</td>
</tr>
<tr>
<td>2</td>
<td>1.948</td>
<td>2.235</td>
<td>2.807</td>
</tr>
<tr>
<td>3</td>
<td>2.113</td>
<td>2.387</td>
<td>2.929</td>
</tr>
<tr>
<td>4</td>
<td>2.226</td>
<td>2.489</td>
<td>3.011</td>
</tr>
<tr>
<td>5</td>
<td>2.312</td>
<td>2.569</td>
<td>3.039</td>
</tr>
</tbody>
</table>

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
Versatile tests for non-linear data

The score vectors and values of $U$'s which are computed from the table are as follows:

\[ a_1 = (-0.097, 0.026, 0.097, 0.125, 0.151), \quad a_2 = (0.047, -0.091, 0.022, 0.108, 0.204), \]
\[ a_3 = (-0.005, 0.022, -0.258, -0.039, 0.423), \quad a_4 = (0.001, -0.006, 0.289, -0.108, 0.288), \]
\[ U_1 = -1.52, \quad U_2 = -1.97, \quad U_3 = 2.11, \quad \text{and} \quad U_4 = 0.27. \]

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

\[ Q_1 = 2.32 \quad (p = 0.128), \quad Q_2 = 6.21 \quad (p = 0.045), \quad Q_3 = 10.67 \quad (p = 0.014) \]

and \( Q_t \)-test detects non-linearity 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

<table>
<thead>
<tr>
<th></th>
<th>CR</th>
<th>PR</th>
<th>MR</th>
<th>NR</th>
<th>PD</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>1</td>
<td>16</td>
<td>5</td>
<td>25</td>
<td>26</td>
<td>72</td>
</tr>
<tr>
<td>Group 2</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>36</td>
<td>28</td>
<td>73</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>21</td>
<td>6</td>
<td>61</td>
<td>54</td>
<td>145</td>
</tr>
</tbody>
</table>

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 sketch it in the next subsection.

3.1. The normal approximation of multiple hypergeometric distribution

When \( Y_1 \) and \( Y_2 \) are independently distributed multinomial random vectors with the parameters \( n_1,(p_{11}, \ldots, p_{1k}) \) and \( n_2,(p_{21}, \ldots, p_{2k}) \), respectively, we have,

\[
Pr[(Y_{21}, \ldots, Y_{2k}) = (y_{21}, \ldots, y_{2k}) | C] = \frac{\psi_{j1}^{y_{j1}} \cdots \psi_{jk}^{y_{jk}} / \prod_{i=1}^{k} \prod_{j=1}^{k} y_{ij}!}{\sum_{r_{21}+\cdots+r_{2k}=n_2} \psi_{j1}^{r_{j1}} \cdots \psi_{jk}^{r_{jk}} / \prod_{i=1}^{k} \prod_{j=1}^{k} r_{ij}!},
\]

(3.1)

where \( \psi_j \) is the odds ratio of \( j \)-th category with respect to category 1, i.e.

\[ \psi_j = p_{1j} p_{2j} / p_{2j} p_{1j}, \quad (j = 1, 2, \ldots, k). \]
Note that $v_i \equiv 1$. 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 $j = 1, 2, \ldots, k$.

**Lemma 3.1.** (Sinkhorn(1967)) If \( \{m_{ij}\} \) satisfy
\[ \sum_{j=1}^{k} m_{ij} = n_i, \text{ for } i = 1, 2, \ldots, N; \]
\[ \sum_{i=1}^{k} m_{ij} = \tau_j, \text{ for } j = 1, 2, \ldots, k; \]
and \( (m_{1j}, m_{2j})/(m_{2i}n_{ij}) = \tau_j \) for $j = 1, 2, \ldots, k$, then \( \{m_{ij}\} \) are uniquely determined by the following algorithm:

\[
\begin{align*}
(m^{(1)}_{1j}) & = \frac{n_1}{k}, j = 1, 2, \ldots, k, \\
(m^{(1)}_{2i}) & = \frac{n_2}{k[1 + \sum_{j=2}^{k} (\psi_j - 1)/k]}, \\
(m^{(1)}_{2j'}) & = \frac{n_2 \psi_j}{k[1 + \sum_{j=2}^{k} (\psi_j - 1)/k]}, j' = 2, \ldots, k, \\
(m^{(2)}_{ij}) & = \frac{m^{(1)}_{ij} \tau_j}{m^{(1)}_{ij}}, \\
(m^{(3)}_{ij}) & = \frac{m^{(2)}_{ij} n_i}{m^{(2)}_{i}}, \\
(m^{(2h-1)}_{ij}) & = \frac{m^{(2h-1)}_{ij} \tau_j}{m^{(2h-1)}_{ij}}, \quad h = 1, 2, \ldots, \infty,
\end{align*}
\]

**Theorem 3.2.** (Plackett(1981)) Suppose (A1), then the conditional distribution of $Y = (Y_{22}, \ldots, Y_{2k})'$ conditioned on $C$ converges in distribution to $N_{k-1}(m_2, \Sigma)$ as $N \to \infty$, where $m_2 = (m_{22}, \ldots, m_{2k})'$ and $\Sigma^{-1} = (\sigma_{ij}),$ and $\sigma_{ij} = m_{1i}^{-1} + m_{2i}^{-1} + (m_{1j}^{-1} + m_{2j}^{-1}) \delta_{ij}, i, j = 2, \ldots, k,$ 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 $N^{-r}c_i = O(1)$ if and only if $N^{-r}c_i$ tends to a constant as $N \to \infty$.

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

\[ N^{-r}c_i = O(1), (i = 1, 2, \ldots, k), \]

where $c_i = c_i^r$, 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 $c_i = \sum_{j=1}^{k} \tau_j + (\tau_i - N)/2$ we get $N^{-1}c_i = O(1)$ easily from (A1). \qed
LEMMA 3.4. If (A1) is satisfied, then

\[ N^{-r}(c_r, a_0) a_{op} = O(1), \quad r = 1, 2, \ldots, k - 1, \quad i = 1, 2, \ldots, k. \]

PROOF. From the definition of \( a_0 \), \( a_{op} = 1/N^{1/2} \) for all \( i \). So by Lemma 3.3 we obtain \( N^{-(r+1/2)}(c_r, a_0) = O(1) \). Hence, the desired result follows. \( \square \)

LEMMA 3.5. If (A1) is satisfied, then

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

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

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

Applying Lemma 3.3 and 3.4, it follows that

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

Suppose that the result is true for \( r = 1, 2, \ldots, m - 1 \). We have

\[
\begin{align*}
d_m &= c_m - \sum_{l=0}^{m-1} (c_m, a_l) a_{li}, \\
&= c_m - (c_m, a_0) a_0 - \sum_{l=1}^{m-1} (c_m, d_l) \frac{d_l}{\|d_l\|^2},
\end{align*}
\]

it follows that \( N^{-m}c_m = O(1) \) from Lemma 3.3, and \( N^{-m}(c_m, a_0) a_0 = O(1) \) from Lemma 3.4. Furthermore, since by the assumption of induction and Lemma 3.4 \( N^{-m-l} (c_m, d_l) = O(1) \) for \( l = 1, 2, \ldots, m - 1 \), and also from the assumption of induction we have \( N^l d_l/\|d_l\|^2 = O(1) \) for \( l = 1, 2, \ldots, m - 1 \). Therefore

\[ N^{-m} (c_m, d_l) \frac{d_l}{\|d_l\|^2} = O(1). \]

So the result is true for \( r = m \). By the induction the proof is completed. \( \square \)

Using these lemmas we may easily show

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

3.3. Proof of Theorem 2.2

LEMMA 3.7. Assume (A1), then under \( H_0 \) the conditional distribution of \( N^{-1/2}(Y - N^{-1/2}(c_r, a_0)) \) conditioned on \( C \) converges in distribution to \( N_{k-1}(0, \Sigma_0) \) as \( N \to \infty \), where

\[ \rho = (\rho_2, \ldots, \rho_k)' \text{ and } \Sigma_0^{-1} = (\sigma_{ij0}) \text{ with } \sigma_{ij0} = [\rho_{i1}^{-1} + \delta_{ij}^{-1}] / \gamma_1 \gamma_2, \quad i, j = 2, \ldots, k. \]
PROOF. Under $H_0$, $m_{ij} = N^{-1/2} \rho_{ij}$, for $i = 1, 2$, and $j = 1, \ldots, k$. Thus from Theorem 3.2 we have the desired result.

Now assuming stronger assumption than (A1)
Assumption (A2) $N^{1/2}(n_j/N - \gamma_j) \to 0$, $(0 < \gamma_j < 1; i = 1, 2)$ and $N^{1/2}(\tau_j/N - \rho_j) \to 0$ $(0 < \rho_j < 1; j = 1, 2, \ldots, k)$ as $N \to \infty$

we prove Theorem 2.2. Put $B = (a_{rj} - a_{r1})$, $r = 1, 2, \ldots, t$, and $j = 2, 3, \ldots, k$. Then from (2.3) and (A2) we may represent

$$U = \frac{AY_2}{\sqrt{n_1n_2/N(N - 1)}}$$

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

$$B\Sigma_0B'/\{n_1n_2/N^2(N - 1)\} = I_t + o(1).$$

3.4. Proof of Theorem 2.3

Recall that the alternative hypothesis we consider is

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

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

$$m^{(1)}_{ij} = \frac{n_2}{k},$$

$$m^{(1)}_{2j} = \frac{n_2}{k} \left[ 1 - \sum_{j=2}^{k} \frac{(\psi_j - 1)}{k} + o(N^{-1/2}) \right],$$

$$m^{(1)}_{jj'} = \frac{n_2}{k} \left[ \psi_{j'} - \sum_{j=2}^{k} \frac{(\psi_j - 1)}{k} + o(N^{-1/2}) \right], \quad j' = 2, 3, \ldots, k,$$

$$m^{(2)}_{ii} = N \gamma_i \rho_1 + (-1)^{i+1} N \gamma_i \gamma_2 \rho_1 \sum_{j=2}^{k} \frac{(\psi_j - 1)}{k} + o(N^{1/2}),$$

$$m^{(2)}_{ij'} = N \gamma_i \rho_j + (-1)^i N \gamma_i \gamma_2 \rho_j \left[ \psi_{j'} - 1 - \sum_{j=2}^{k} \frac{\psi_j - 1}{k} \right] + o(N^{1/2}),$$

$$m^{(1)}_{ij} = N \gamma_i \rho_j + N^{1/2} \eta_{ij} + o(N^{1/2}),$$
for \( l = 3, 4, \cdots \), where

\[
\eta_{11} = (-1)^{i+1}X^{1/2}g_{11}^2\rho_{1} \sum_{j=2}^{k}(\psi_{j} - 1)\rho_{j},
\]

\[
\eta_{1j'} = (-1)^{i}X^{1/2}g_{11}^2\rho_{j'}\left[\psi_{j'} - 1 - \sum_{j=2}^{k}(\psi_{j} - 1)\rho_{j}\right], \quad (j' = 2, 3, \cdots, k).
\]

**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 \( m_{1j}^{(1)}, m_{21}^{(1)}, m_{2j}^{(1)}, m_{11}^{(2)} \) and \( m_{1j}^{(2)} \). Similarly substituting \( \psi_{j} \) 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 \( N^{-1/2}(Y - N\gamma_{2}\rho) \) conditioned on \( C \) converges in distribution to \( N_{k-1}(\eta_{2}, \Sigma_{0}) \) as \( N \to \infty \) under \( H_{1} \), where \( \eta_{2} = (\eta_{22}, \cdots, \eta_{2k})' \). Now since

\[
U = (N^{1/2}B)N^{-1/2}(Y - N\gamma_{2}\rho)/\sqrt{n_{1}n_{2}/N(N - 1) + o(1)},
\]

the conditional distribution of \( U \) conditioned on \( C \) converged in distribution to \( N_{t}(B\eta_{2} \sqrt{n_{1}n_{2}/N^{2}(N - 1)} , B\Sigma_{0}B'/(n_{1}n_{2}/N^{2}(N - 1))) \). From the remark at the end of the previous subsection we have

\[
\lim_{N \to \infty} B\Sigma_{0}B'/(n_{1}n_{2}/N^{2}(N - 1)) \to I_{t}.
\]

Furthermore, since

\[
\sum_{j=1}^{k}a_{rj}\rho_{j} = 0, \quad \text{and} \quad \sum_{j=1}^{k}\rho_{j} = 1
\]

the \( r \)-th element of \( B\eta_{2} \) is simply represented by

\[
\delta_{r} = \sum_{j=1}^{k}(a_{rj} - a_{r1})\eta_{2j} = N^{1/2}g_{11}^2\sum_{j=1}^{k}a_{rj}\rho_{j}(\psi_{j} - 1).
\]

This completes the proof of Theorem 2.3.

**References**


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.


Received October 3, 1998

Revised November 19, 1999