TESTING ORDERED CATEGORICAL DATA IN $ 2 ¥times k $ TABLES BY THE STATISTICS WITH ORTHONORMAL SCORE VECTORS

Jayasekara, Leslie
Graduate School of Mathematics, Kyushu University

Yanagawa, Takashi
Graduate School of Mathematics, Kyushu University

https://doi.org/10.5109/13447

出版情報：Bulletin of informatics and cybernetics. 27 (1), pp.129-145, 1995-03. Research Association of Statistical Sciences

Bulletin of Informatics and Cybernetics, Vol. 27, No. 1, 1995

TESTING ORDERED CATEGORICAL DATA IN $2 \times k$ TABLES BY THE STATISTICS WITH ORTHONORMAL SCORE VECTORS

By

Leslie JAYASEKARA* and Takashi YANAGAWA*

Abstract

A class of tests is proposed for detecting the difference of two populations in an ordinal categorical table. Characteristics of the proposed tests are studied. It will be shown that the new tests may have higher powers for a class of non-linear responses than the other conventional tests.

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

To fix the idea consider a randomized clinical trial for testing the effectiveness of a new drug against a placebo. Frequently the effectiveness of the drug is measured by ordinal categories such as remarkable (+++++), effective (++++), $\cdots$, slightly effective (+), and not effective. Table 1 displays data from such a trial. The data is plotted in figure 1, except for the data in not effective category. The figure shows a $\bigcup$ shaped response. Normaly treatment by drugs may change the environment of patient's interior such as immune system and we often come across $\bigcup$ shaped, or more complicated responses. The purpose of the present paper is to propose a class of statistical tests which have higher powers for those non-linear responses in a $2 \times k$ ordered categorical tables.

Table 1: 2×6 table with ordered categories

	Effectiveness						
Drugs	Not effective	+	++	+++	++++	+++++	Total
Placebo	65	3	3	3	3	3	80
Treatment	56	8	3	2	3	8	80

The Wilcoxon test [6], or equivalently Mantel's extended test [3] has been applied for testing ordered categorical data in $2 \times k$ tables. The test has no high powers for

* Graduate School of Mathematics, Kyushu University 33,Fukuoka 812-81, Japan.

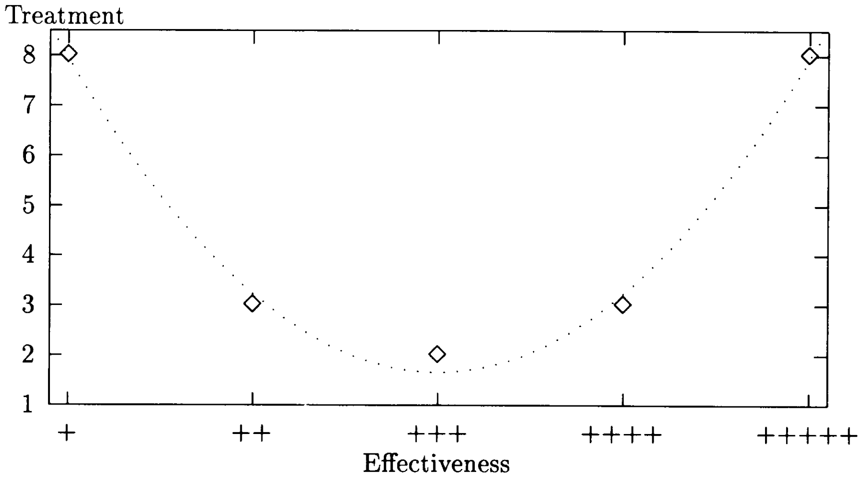

Figure 1: Plotted treatment group in Table 1.

the responses we are interested in. Alternatively, the cumulative chi-square test [5] and Nair's test [4] may be applied. The former test is an omnibus test developed for a wider class of alternatives and the latter test was, in particular, designed to detect the dispersion alternatives. Jayasekara, Yanagawa and Tsujitani [1] have developed a test which is useful for detecting location-dispersion alternatives. In this paper we generalize the location-dispersion test by constructing the statistic whose score vectors are orthonormal.

In section 2 we propose the test statistic. Its asymptotic distributions under the null hypothesis and contiguous alternatives are studied in section 3. We compare the powers of the tests with the competitors in section 4.

2. The Test Statistic

Consider $2 \times k$ table given in Table 2, and suppose that $X = (X_1, X_2, \cdots, X_k)$ and $Y = (Y_1, Y_2, \cdots, Y_k)$ are independently distributed multinomial random vectors. We consider the following null hypothesis:

H_0 : X and Y are identically distributed.

To define the test statistic for H_0, the orthonormal scores will be introduced based on the Wilcoxon score.

Let c_i be the Wilcoxon Score defined by $c_i = \sum_{j=1}^{i-1} \tau_j + (\tau_i - N)/2$ for $i = 1, 2, \cdots, k$, so that $\sum_{i=1}^{k} \tau_i c_i = 0$. We define the inner product of $\mathbf{a} = (a_1, \cdots, a_k)$ and $\mathbf{b} = (b_1, \cdots, b_k)$ by $(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{k} \tau_i a_i b_i$ and also $\|\mathbf{a}\|^2 = (\mathbf{a}, \mathbf{a})$.

Let c_j^i be the i-th power of c_j, $j = 1, 2, \cdots, k$ and put $\mathbf{c}_i = (c_1^i, c_2^i, \cdots, c_k^i)$, $i = 0, 1, \cdots, k$. It is obvious that $\mathbf{c}_0, \mathbf{c}_1, \cdots, \mathbf{c}_k$ are linearly independent. Let $\mathbf{a}_0, \mathbf{a}_1, \cdots, \mathbf{a}_k$ be orthonormal score vectors which are obtained by appling Gram-Schmidit orthonor-

Table 2: $2 \times k$ contingency table

	Ordered Categories						Total
X	X_1	X_2	.	.	.	X_k	n_1
Y	Y_1	Y_2	.	.	.	Y_k	n_2
Total	τ_1	τ_2	.	.	.	τ_k	N

malization to these vectors. That is, $\mathbf{a}_0 = \frac{\mathbf{c}_0}{\|\mathbf{c}_0\|}$ and $\mathbf{a}_r = \frac{\mathbf{d}_r}{\|\mathbf{d}_r\|}$, $r = 1, 2, \cdots, k$, where $\mathbf{d}_r = \mathbf{c}_r - \sum_{l=0}^{r-1}(\mathbf{c}_r, \mathbf{a}_l)\mathbf{a}_l$. We have,

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

Putting $\mathbf{a}_r = (a_{r1}, a_{r2}, \cdots, a_{rk})$, and

$$S_r = \sum_{i=1}^{k} a_{ri} Y_i, \quad \text{for } r = 1, 2, \cdots, k,$$

the test statistic which we propose in this paper for testing the null hypothesis H_0 is given by

$$Q_t = \sum_{r=1}^{t} \frac{S_r^2}{V_0[S_r|C]}, \quad \text{for each } t \in \{1, 2, \cdots, k\},$$

where $V_0[\cdot|C]$ is the conditional variance conditioned on $C = \{n_1, n_2, \tau_1, \cdots, \tau_k\}$ under H_0 and is given by

$$V_0[S_r|C] = \frac{n_1 n_2}{N(N-1)}, \quad \text{for } r = 1, 2, \cdots, k.$$

Note that the conditional distribution of Y under H_0 conditioned on C is given by

$$\Pr[(Y_1, \cdots, Y_k) = (y_1, \cdots, y_k)|C] = \frac{\binom{\tau_1}{y_1} \cdots \binom{\tau_k}{y_k}}{\binom{N}{n_2}}.$$

It is shown in section 3 that Q_t follows a chi-square distribution with t degrees of freedom under H_0. This test statistic Q_t is identical to the Wilcoxon test statistic and the location-dispersion test statistic for $t = 1$ and 2, respectively.

Investigating the asymptotic powers and also by simulation, it is shown that Q_3 has the highest power for the $\cap$ shaped or $\cup$ shaped response among $\{Q_t\}$; and Q_4 has the highest power for the $\cap\!\cup$ shaped or $\cup\!\cap$ shaped response among $\{Q_t\}$.

3. Asymptotic Distributions

First we approximate a multiple non-central hypergeometric distribution by the binomial distribution. Then using this approximation, we develop the asymptotic distributions of Q_t under the hypothesis and also under contiguous alternatives.

3.1. The Multiple Non-Central Hypergeometric Distribution

When X and Y are independently distributed multinomial random variables with the parameters $n_1, \pi_1 = (\pi_{11}, \cdots, \pi_{1k})$ and $n_2, \pi_2 = (\pi_{21}, \cdots, \pi_{2k})$, respectively, we have,

$$\Pr[(Y_2, \cdots, Y_k) = (y_2, \cdots, y_k)|C] = \frac{g(\mathbf{y})\psi_1^{y_1} \cdots \psi_k^{y_k}}{\sum_{j_1+\cdots+j_k=n_2} g(\mathbf{j})\psi_1^{j_1} \cdots \psi_k^{j_k}}, \tag{2}$$

where $g(\mathbf{y}) = n_1!n_2!/[y_1! \cdots y_k!(\tau_1 - y_1)! \cdots (\tau_k - y_k)!]$ and $\psi_j = \pi_{11}\pi_{2j}/\pi_{21}\pi_{1j}$ which is the odds-ratio parameter, relative to category 1, for $j = 1, 2, \cdots, k$, so that $\psi_1 \equiv 1$. We use the following assumptions for τ_j and n_2 which are defined in Table 2.

(A1) $\lim_{N\to\infty} \frac{\tau_j}{N} = 0$, for $j = 2, 3, \cdots, k$.

(A2) $\frac{n_2}{N} = p$ for some given p such that $0 < p < 1$.

THEOREM 3.1. *Suppose that* (A1) *and* (A2) *are satisfied, then*

$$\lim_{N\to\infty} \frac{g(\mathbf{y})\tau_1! \cdots \tau_k!}{N! \prod_{j=2}^k \binom{\tau_j}{y_j} p^{y_j}(1-p)^{\tau_j - y_j}} = 1.$$

PROOF. Since

$$g(\mathbf{y})\frac{\tau_1! \cdots \tau_k!}{N!} = \frac{\frac{n_2!}{(n_2-y_2-\cdots-y_k)!}\frac{n_1!}{(n_1-\tau_2-\cdots-\tau_k+y_2+\cdots+y_k)!}}{\frac{N!}{(N-\tau_2-\cdots-\tau_k)!}} \binom{\tau_2}{y_2} \cdots \binom{\tau_k}{y_k},$$

from (A1) and (A2) we have

$$\lim_{N\to\infty} g(\mathbf{y})\frac{\tau_1! \cdots \tau_k!}{N!} = \prod_{j=2}^k \binom{\tau_j}{y_j} p^{y_j}(1-p)^{\tau_j - y_j}.$$

This completes the proof of the theorem. □

From the theorem it follows that the numerator of (2) is approximated by

$$\prod_{j=2}^k \binom{\tau_j}{y_j} (p\psi_j)^{y_j}(1-p)^{\tau_j - y_j} \frac{N!}{\tau_1! \cdots \tau_k!}.$$

Therefore by normalizing this we may approximate the distribution (2) by

$$\Pr[(Y_2, \cdots, Y_k) = (y_2, \cdots, y_k)|C]$$

$$\approx \prod_{j=2}^{k} \binom{\tau_j}{y_j} \left(\frac{p\psi_j}{p\psi_j + 1 - p}\right)^{y_j} \left(\frac{1-p}{p\psi_j + 1 - p}\right)^{\tau_j - y_j} \quad (3)$$

Thus according to this approximation Y_j is binomialy distributed with parameters τ_j and $p\psi_j/(p\psi_j + 1 - p)$, and $Y_2, Y_3, \cdots, Y_k$ are independent.

3.2. Asymptotic Distribution of Q_t Under H_0

Now we consider an approximation of $\Pr[(Y_2, \cdots, Y_k) = (y_2, \cdots, y_k)|C]$ for large τ_j, $j = 2, 3, \cdots, k$. To begin with we shall make the following assumption.

(A3) $N^{-\epsilon}\tau_j = O(1)$, $j = 2, 3, \cdots, k$, for some ϵ such that $0 < \epsilon < \frac{1}{2}$.

The notation $N^{-\epsilon}\tau_j = O(1)$ which is used in this paper, means $N^{-\epsilon}\tau_j$ tends to a constant as $N \to \infty$. Note that (A3) includes (A1). We need the following Lemmas 3.2, 3.3, 3.4 and 3.5 to get Lemma 3.6.

LEMMA 3.2. *If* (A3) *is satisfied, then*

$$N^{-r\epsilon}c_{r1} = O(1) \quad and \quad N^{-r}c_{ri} = O(1), \quad (i = 2, 3, \cdots, k)$$

where $c_{ri} = c_i^r$, *is the r-th power of the i-th Wilcoxon score, for* $r = 1, 2, \cdots, k$.

PROOF. By the definition of c_1 we have $c_1 = -\frac{1}{2}\sum_{i=2}^{k}\tau_i$, and from (A3) it follows that $N^{-\epsilon}c_1 = O(1)$. By the definition of c_i, for $i = 2, 3, \cdots, k$, we have $c_i = \tau_1 + \sum_{j=2}^{k}\tau_j + (\tau_i - N)/2$. ¿From (A3) we can get $N^{-1}c_i = O(1)$ since $N^{-1}\tau_1 = O(1)$. Thus, the proof is completed. □

LEMMA 3.3. *If* (A3) *is satisfied, then*

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

PROOF. ¿From the definition of $\mathbf{a}_0$ we know that $a_{0i} = 1/N^{1/2}$ for all i. So by Lemma 3.2 we obtain $N^{-r-\epsilon+1/2}(\mathbf{c}_r, \mathbf{a}_0) = O(1)$. Hence, the desired result follows. □

LEMMA 3.4. *If* $N^{-(l-1)-\epsilon}d_{l1} = O(1)$, *and* $N^{-l}d_{li} = O(1)$, *for* $i = 2, \cdots, k$, *and* (A3) *is satisfied, then*

(i) $N^{-2l-\epsilon}\|\mathbf{d}_l\|^2 = O(1)$,

(ii) $N^{-(r-1)-\epsilon}(\mathbf{c}_r, \mathbf{d}_l)\frac{d_{l1}}{\|\mathbf{d}_l\|^2} = O(1)$ *and* $N^{-r}(\mathbf{c}_r, \mathbf{d}_l)\frac{d_{li}}{\|\mathbf{d}_l\|^2} = O(1)$, $i = 2, 3, \cdots, k$.

PROOF. (i) The result can be obtained by the definition of $\mathbf{d}_l$.

(ii) Expanding $(\mathbf{c}_r, \mathbf{d}_l)$ and applying Lemma 3.2 we can obtain $N^{r+l+\epsilon}(\mathbf{c}_r, \mathbf{d}_l) = O(1)$, for all r, l. Then using (i), the result follows.

□

For the assumption of Lemma 3.4 we have

LEMMA 3.5. *If* (A3) *is satisfied, then*

$$N^{-(r-1)-\epsilon}d_{r1} = O(1) \ \ and \ \ N^{-r}d_{ri} = O(1), \ for\ r = 1, 2, \cdots, k \ and\ i = 2, 3, \cdots, k.$$

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

$$d_{11} = c_{1i} - (\mathbf{c}_1, \mathbf{a}_0)a_{0i}, \qquad \text{for } i = 1, 2, \cdots, k.$$

Applying Lemma 3.2 and 3.3, it follows that

$$N^{-\epsilon}d_{11} = O(1) \ \ and \ \ N^{-1}d_{1i} = O(1), \qquad \text{for } i = 2, 3, \cdots, k.$$

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

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

it follows that $N^{-(m-1)-\epsilon}d_{m1} = O(1)$ and $N^{-m}d_{mi} = O(1)$ from Lemma 3.2, 3.3 and 3.4. So the result is true for $r = m$. By the induction the proof is completed. □

Using these lemmas we may show,

LEMMA 3.6. *If* (A3) *is satisfied then*

$$N^{1-\epsilon/2}a_{r1} = O(1) \ \ and \ \ N^{\epsilon/2}a_{ri} = O(1),$$

for $r = 1, 2, \cdots, k$ *and* $i = 2, 3, \cdots, k$.

PROOF. It is straightforward to show the lemma from the definition of $\mathbf{a}_r$, $r = 1, 2, \cdots, k$, and by Lemma 3.5. □

Denote by $E_{A0}[\cdot]$ and $V_{A0}[\cdot]$ the expectation and variance, respectively, under H_0 when Y_j, $j = 2, 3, \cdots, k$, follow distribution given in (3), thus we have, for example,

$$\begin{aligned} E_{A0}[Y_j] &= \tau_j p, & (4) \\ V_{A0}[Y_j] &= \tau_j p(1-p). & (5) \end{aligned}$$

We have the following lemma.

LEMMA 3.7. *If* (A2) *and* (A3) *are satisfied, then*

(i) $\mathrm{E}_{A0}[S_r] = 0$, $r = 1, 2, \cdots, k$.

(ii) $\lim_{N\to\infty} \frac{V_0[S_r|C]}{V_{A0}[S_r]} = 1$, $r = 1, 2, \cdots, k$.

PROOF. Since $\sum_{i=1}^{k} Y_i = n_2$,

$$S_r = a_{r1}n_2 + \sum_{i=2}^{k}(a_{ri} - a_{r1})Y_i. \tag{6}$$

Thus from (4)

$$\begin{aligned} E_{A0}[S_r] &= a_{r1}n_2 + \sum_{i=2}^{k}(a_{ri} - a_{r1})\tau_i p, \\ &= a_{r1}n_2 + N^{1/2}(\mathbf{a}_0, \mathbf{a}_r)p - a_{r1}Np. \end{aligned}$$

Furthermore, from (1), we have for $r = 1, 2, \cdots, k$,

$$E_{A0}[S_r] = a_{r1}N\left(\frac{n_2}{N} - p\right).$$

Thus from (A2) we have (i). Next we prove (ii). ¿From (5) and (6) we have

$$\begin{aligned} V_{A0}[S_r] &= \sum_{i=2}^{k}(a_{ri} - a_{r1})^2\tau_i p(1-p) \\ &= p(1-p)(1 + a_{r1}^2 N). \end{aligned}$$

Therefore using (A2) and Lemma 3.6 we have

$$\frac{V_0[S_r|C]}{V_{A0}[S_r]} = \frac{n_1n_2/N(N-1)}{p(1-p)(1+a_{r1}^2 N)} \to 1 \quad as \quad N \to \infty.$$

□

THEOREM 3.8. *If* (A2) *and* (A3) *are satisfied, then under* H_0, *the conditional distributions of* $S_r/\sqrt{V_0[S_r|C]}$, $r = 1, 2, \cdots, k$, *given* C *is approximated by a standard normal distribution for large* N.

PROOF. We can write

$$\Pr\left[\frac{S_r}{\sqrt{V_0[S_r|C]}} \le x|C\right] = \Pr\left[\frac{S_r - E_{A0}[S_r]}{\sqrt{V_{A0}[S_r]}} \le \sqrt{\frac{V_0[S_r|C]}{V_{A0}[S_r]}}\left\{x - \frac{E_{A0}[S_r]}{\sqrt{V_0[S_r|C]}}\right\}|C\right].$$

When $N \to \infty$, (A3) $\Rightarrow$ (A1), and the distribution of $(Y_2, Y_3, \cdots, Y_k)$ may be approximated by the multiple of the independent binomial distributions. Furthermore $\tau_j \to \infty$, $j = 2, 3, \cdots, k$ when $N \to \infty$ from (A3). Therefore the distribution of $(S_r - E_{A0}[S_r])/\sqrt{V_{A0}[S_r]}$ may be approximated by $N(0,1)$ for large N. Then the result follows from Lemma 3.7 and by Slutsky theorem (See Lehmann [2] Appendix, Sec. 3, Corollary 2). □

THEOREM 3.9. *If* (A2) *and* (A3) *are satisfied then the conditional distribution of* Q_t *given* C *under* H_0 *is approximated by a chi-squared distribution with* t *degrees of freedom for large* N.

PROOF. It is sufficient to show that the conditional covariance of S_{r_1} and S_{r_2} conditioned on C under H_0, which is denoted by $Cov_0[S_{r_1}, S_{r_2}|C]$, is 0. ¿From (6) we have

$$\begin{aligned} Cov_0[S_{r_1}, S_{r_2}|C] &= \sum_{i=2}^{k}(a_{r_1 i} - a_{r_1 1})(a_{r_2 i} - a_{r_2 1})V_0[Y_i|C] \\ &\quad + \sum_{i=2}^{k}\sum_{j=2_{i\neq j}}^{k} (a_{r_1 i} - a_{r_1 1})(a_{r_2 j} - a_{r_2 1})Cov_0[Y_i, Y_j|C] \\ &= \sum_{i=2}^{k}(a_{r_1 i} - a_{r_1 1})(a_{r_2 i} - a_{r_2 1})\frac{\tau_i(N-\tau_i)n_1 n_2}{N^2(N-1)} \\ &\quad - \sum_{i=2}^{k}\sum_{j=2_{i\neq j}}^{k} (a_{r_1 i} - a_{r_1 1})(a_{r_2 j} - a_{r_2 1})\frac{\tau_i\tau_j n_1 n_2}{N^2(N-1)} \\ &= \sum_{i=2}^{k}(a_{r_1 i} - a_{r_1 1})(a_{r_2 i} - a_{r_2 1})\frac{\tau_i n_1 n_2}{N(N-1)} \\ &\quad - \sum_{i=2}^{k}\sum_{j=2}^{k}(a_{r_1 i} - a_{r_1 1})(a_{r_2 j} - a_{r_2 1})\frac{\tau_i\tau_j n_1 n_2}{N^2(N-1)} \end{aligned}$$

Now

$$\sum_{i=2}^{k}(a_{r_1 i} - a_{r_1 1})(a_{r_2 i} - a_{r_2 1})\tau_i = a_{r_1 1}a_{r_2 1}N$$

$$\begin{aligned} \sum_{i=2}^{k}\sum_{j=2}^{k}(a_{r_1 i} - a_{r_1 1})(a_{r_2 j} - a_{r_2 1})\tau_i\tau_j &= \left[\sum_{i=2}^{k}(a_{r_1 i} - a_{r_1 1})\tau_i\right]\left[\sum_{j=2}^{k}(a_{r_2 j} - a_{r_2 1})\tau_j\right] \\ &= a_{r_1 1}a_{r_2 1}N^2. \end{aligned}$$

Thus substituting these equalities into the above formula, we have the desired result. □

3.3. Asymptotic Distribution of Q_t Under the Alternative

In this section we obtain asymptotic distribution of Q_t under the alternative hypothesis $H_1 : \psi_j = 1 + A_j/N^{\epsilon/2}$, $j = 1, 2, \cdots, k$, where A_j is a constant and $0 < \epsilon < 1/2$. We denote by $E_A[\cdot]$ and $V_A[\cdot]$ the expectation and variance, respectively, under H_1 when Y_j, $j = 2, 3, \cdots, k$, follow the distribution given in (3). Thus for example,

$$E_A[Y_j] = \frac{\tau_j p\psi_j}{p\psi_j + 1 - p}, \tag{7}$$

$$V_A[Y_j] = \frac{\tau_j p(1-p)\psi_j}{(p\psi_j + 1 - p)^2}. \tag{8}$$

LEMMA 3.10. *If* (A2) *and* (A3) *are satisfied, then*

(i) $\lim_{N\to\infty} \frac{E_A[S_r]}{\sqrt{V_0[S_r|C]}} = l_{S_r}$, *where* $l_{S_r} = \sqrt{p(1-p)}\sum_{j=2}^{k}(a_{rj}-a_{r1})\tau_j(\psi_j-1)$.

(ii) $\lim_{N\to\infty} \frac{V_0[S_r|C]}{V_A[S_r]} = 1,\ r = 1,2,\cdots,k.$

PROOF. (i) From (6) and (7) we have

$$E_A[S_r] \quad = \quad a_{r1}n_2 + \sum_{j=2}^{k}(a_{rj}-a_{r1})\frac{\tau_j p\psi_j}{p\psi_j+1-p}.$$

Now since

$$E_{A0}[S_r] = a_{r1}n_2 + \sum_{j=2}^{k}(a_{rj}-a_{r1})\tau_j p,$$

and furthermore $E_{A0}[S_r] = 0$, we have

$$\begin{aligned} E_A[S_r] &= (E_A[S_r]-E_{A0}[S_r]) + E_{A0}[S_r] \\ &= \sum_{j=2}^{k}(a_{rj}-a_{r1})\tau_j\left[\frac{p(1-p)(\psi_j-1)}{p\psi_j+1-p}\right]. \end{aligned}$$

Thus

$$\frac{E_A[S_r]}{\sqrt{V_0[S_r|C]}} = \sqrt{\frac{N(N-1)}{n_1n_2}}\sum_{j=2}^{k}(a_{rj}-a_{r1})\tau_j\frac{p(1-p)(\psi_j-1)}{p\psi_j+1-p}.$$

Since

$$\sqrt{\frac{N(N-1)}{n_1n_2}}p(1-p) \to \sqrt{p(1-p)}, \qquad \text{(from (A2))}$$

$(p\psi_j+1-p) \to 1$ *as* $N\to\infty$, and $(a_{rj}-a_{r1})\tau_j(\psi_j-1) = O(1)$, the result follows.
(ii) From (6) and (8) we have

$$V_A[S_r] \quad = \quad \sum_{j=2}^{k}(a_{rj}-a_{r1})^2\frac{\tau_j p(1-p)\psi_j}{(p\psi_j+1-p)^2}.$$

Thus it follows that

$$\frac{V_0[S_r|C]}{V_A[S_r]} = \frac{n_1n_2/N(N-1)}{p(1-p)A},$$

where $A = \sum_{j=2}^{k}(a_{rj}-a_{r1})^2\tau_j\psi_j/(p\psi_j+1-p)^2$.
Under H_1,

$$\begin{aligned} A &= \sum_{j=2}^{k}(a_{rj}-a_{r1})^2\tau_j\left(1+\frac{A_j}{N^{\epsilon/2}}\right)\left(1+p\frac{A_j}{N^{\epsilon/2}}\right)^{-2} \\ &= \sum_{j=2}^{k}(a_{rj}-a_{r1})^2\tau_j\left[1+(1-2p)\frac{A_j}{N^{\epsilon/2}}+p(3p-2)\frac{A_j^2}{N^{\epsilon}}+o\left(\frac{1}{N^{\epsilon}}\right)\right]. \end{aligned}$$

Furthermore since

$$\sum_{j=2}^{k}(a_{rj}-a_{r1})^2\tau_j = 1 + a_{r1}^2 N,$$

we have

$$A = 1 + a_{r1}^2 N + \sum_{j=2}^{k}(a_{rj}-a_{r1})^2\tau_j \left[(1-2p)\frac{A_j}{N^{\epsilon/2}} + p(3p-2)\frac{A_j^2}{N^{\epsilon}} + o\left(\frac{1}{N^{\epsilon}}\right)\right].$$

Employing Lemma 3.6 and (A3) we may show that $a_{r1}^2 N \to 0$ and

$$\sum_{j=2}^{k}(a_{rj}-a_{r1})^2\tau_j \left[(1-2p)\frac{A_j}{N^{\epsilon/2}} + p(3p-2)\frac{A_j^2}{N^{\epsilon}} + o\left(\frac{1}{N^{\epsilon}}\right)\right] \to 0 \;\; as \; N \to \infty.$$

Thus $A \to 1$ *as* $N \to \infty$. Since $n_1 n_2/N(N-1) \to p(1-p)$ *as* $N \to \infty$ (from (A2)), the result follows. □

THEOREM 3.11. *If* (A2) *and* (A3) *are satisfied, then under* H_1 *the conditional distribution of* $S_r/\sqrt{V_0[S_r|C]}$, $r = 1, 2, \cdots, k$, *given* C *is approximated by* $N(l_{S_r}, 1)$ *for large* N, *where* l_{S_r} *is given in Lemma 3.10.*

PROOF. Similarly as the proof of Theorem 3.8, the distribution of $(S_r - E_A[S_r])/\sqrt{V_A[S_r]}$ may be approximated by $N(0,1)$, for large N, under (A3) and

$$\Pr\left[\frac{S_r}{\sqrt{V_0[S_r|C]}} \le x|C\right] = \Pr\left[\frac{S_r - E_A[S_r]}{\sqrt{V_A[S_r]}} \le \sqrt{\frac{V_0[S_r|C]}{V_A[S_r]}}\left\{x - \frac{E_A[S_r]}{\sqrt{V_0[S_r|C]}}\right\}|C\right].$$

Thus it is straightforward to show the theorem from Lemma 3.10 and by Slutsky theorem. □

LEMMA 3.12. *Suppose that* (A2) *and* (A3) *are satisfied, then under* H_1 *the conditional covariance of* S_{r_1} *and* S_{r_2} *given* C *tends to 0 as* $N \to \infty$, *where* $r_1, r_2 = 1, 2, \cdots, k$, *and* $r_1 \ne r_2$.

PROOF. We approximate the conditional distribution of Y_j, $j = 2, 3, \cdots, k$, by the distribution given in (3) and denote by $Cov_A[S_{r_1}, S_{r_2}|C]$ the covariance of S_{r_1} and S_{r_2} under H_1. Then from (8)

$$Cov_A[S_{r_1}, S_{r_2}|C] = \sum_{i=2}^{k}(a_{r_1 i} - a_{r_1 1})(a_{r_2 i} - a_{r_2 1})\tau_i q_i(1-q_i)$$

where $q_i = p\psi_i/(p\psi_i + 1 - p)$ and $\psi_i = 1 + A_i/N^{\epsilon/2}$. Thus

$$|Cov_A[S_{r_1}, S_{r_2}|C]| \le (1/4)\sum_{i=2}^{k}(a_{r_1 i} - a_{r_1 1})(a_{r_2 i} - a_{r_2 1})\tau_i.$$

Furthermore from Lemma 3.6 and (A3) we may show that

$$\sum_{i=2}^{k}(a_{r_1 i} - a_{r_1 1})(a_{r_2 i} - a_{r_2 1})\tau_i = a_{r_1 1}a_{r_2 1}N = \frac{\alpha}{N^{1-\epsilon}}$$

for some constant α. Thus we have $Cov_A[S_{r_1}, S_{r_2}|C] \to 0$ as $N \to \infty$. Therefore we have the desired result. □

¿From Theorem 3.11 and Lemma 3.12 we have the following theorem.

THEOREM 3.13. *If* (A2) *and* (A3) *are satisfied, then the conditional distribution of* Q_t *given* C *under* H_1 *is approximated by a noncentral chi-squared distribution with* t *degrees of freedom and noncentrality parameter* $\sum_{r=1}^{t} l_{S_r}^2$, *when* N *is large.*

4. Power Comparisons

First the asymptotic powers of the test are compared when t is varied, considering several types of non-linear responses. Next the proposed test is compared with the cumulative chi-squared test (CCS test) by simulation.

4.1. Comparison of the Asymptotic Powers when t is Varied

The asymptotic power of the Q_t test under the alternative hypothesis H_1 is given by $\Pr[\chi'^2_t > \chi^2_{t,\alpha}]$, where $\chi^2_{t,\alpha}$ is the upper 100α percentage point of the chi-squared distribution with t degrees of freedom and χ'^2_t is the random variable which follows the density function of the noncentral chi-squared distribution with t degrees of freedom and the noncentrality parameter $\lambda = \sum_{r=1}^{t} l_{S_r}^2$.

Table 3 displays six 2 × 6 tables. Note that Table 3D is identical to Table 1. The data in the treatment groups except for the data belonging to not effective category have been plotted in Figure 2 to visualize the pattern of response. The asymptotic powers of Q_t test when $t = 1, 2, 3, 4$ are computed from each table in Table 3, and shown in Table 4. The table shows that among $\{Q_t\}$, $t = 1, 2, 3, 4$, the test with $t = 2$ provides the maximum powers for the configuration of responses in Table 3A and 3B; for the responses in Table 3C and 3D the test with $t = 3$ provides the maximum powers; and for the responses in Table 3E and 3F the test with $t = 4$ provides the maximum powers.

4.2. Comparison with the CCS test

The simulation studies were conducted to assess the Type I error of the proposed test when $t = 1, 2, 3, 4$ and the CCS test at the significance level $\alpha = 0.05$. We generated 10,000 experiments for each combination of the response probabilities exhibited in the left part of Table 5 and obtained the empirical significance levels. The results are shown in the right part of Table 5. The table shows that when the cell probabilities of the first

Table 3: 2 × 6 tables with ordered categories.

A.

Drugs	Not effective	+ L-1	++ L-2	+++ L-3	++++ L-4	+++++ L-5	Total
		Effectiveness					
Placebo	65	3	3	3	3	3	80
Treatment	56	2	3	5	6	8	80

B.

Drugs	Not effective	+	++	+++	++++	+++++	Total
		Effectiveness					
Placebo	65	3	3	3	3	3	80
Treatment	56	8	6	5	3	2	80

C.

Drugs	Not effective	+	++	+++	++++	+++++	Total
		Effectiveness					
Placebo	65	3	3	3	3	3	80
Treatment	56	2	6	8	6	2	80

D.

Drugs	Not effective	+	++	+++	++++	+++++	Total
		Effectiveness					
Placebo	65	3	3	3	3	3	80
Treatment	56	8	3	2	3	8	80

E.

Drugs	Not effective	+	++	+++	++++	+++++	Total
		Effectiveness					
Placebo	65	3	3	3	3	3	80
Treatment	56	3	8	6	1	6	80

F.

Drugs	Not effective	+	++	+++	++++	+++++	Total
		Effectiveness					
Placebo	65	3	3	3	3	3	80
Treatment	56	6	1	6	8	3	80

Table 4: Asymptotic powers under H_1 ($\alpha = 0.05$).

t		1	2	3	4
Table	3 A	0.465	**0.481**	0.421	0.386
	3 B	0.484	**0.505**	0.442	0.397
	3 C	0.427	0.356	**0.466**	0.428
	3 D	0.416	0.33	**0.424**	0.381
	3 E	0.306	0.351	0.304	**0.415**
	3 F	0.39	0.323	0.317	**0.395**

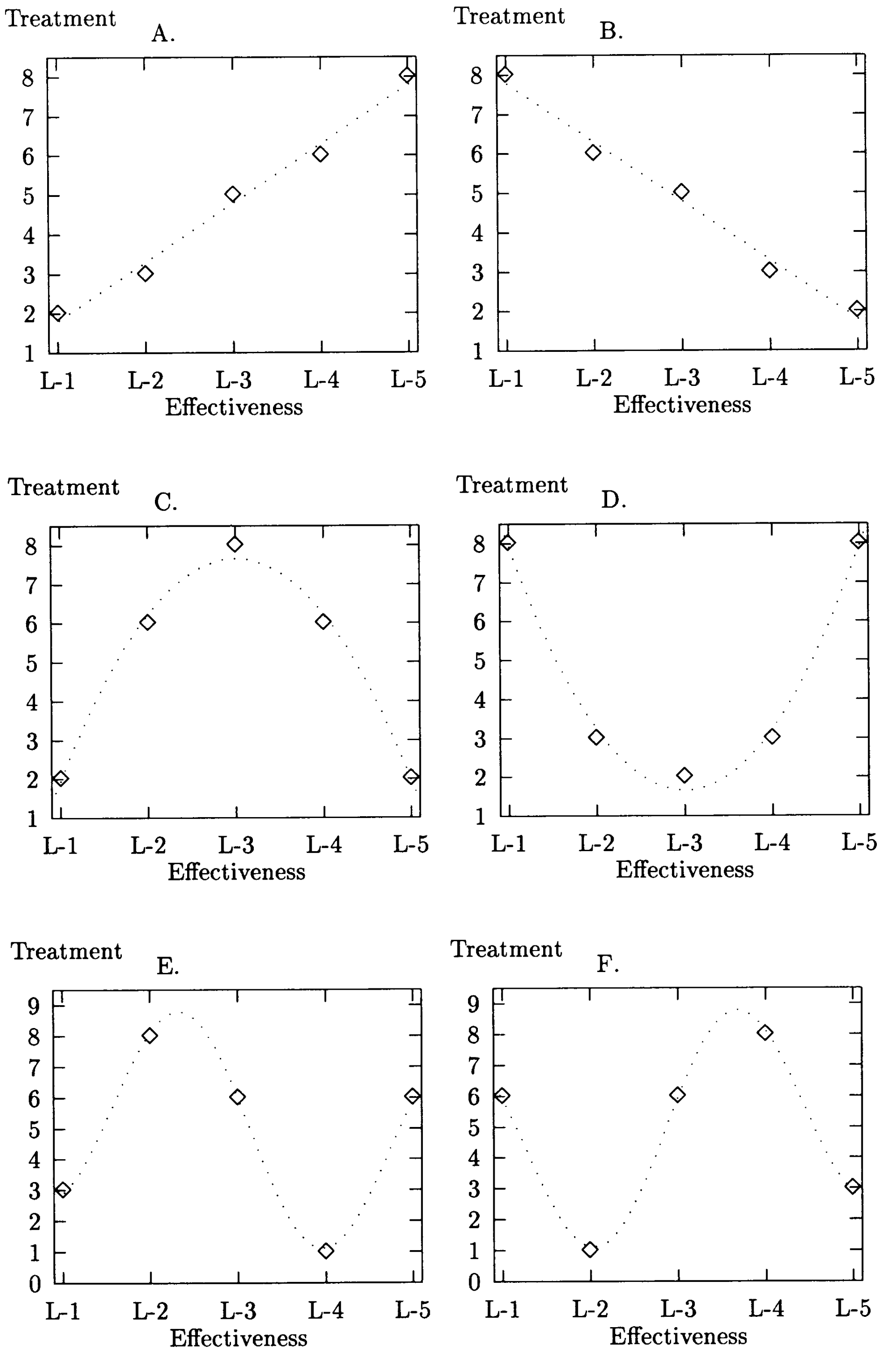

Figure 2: Plotted treatment groups in Table 3.

Table 5: Type I error levels ($\alpha = 0.05$).

Cell probabilities of $(X_1, X_2, \cdots, X_6)$ = Cell probabilities of $(Y_1, Y_2, \cdots, Y_6)$	Type I error				
	Q_1	Q_2	Q_3	Q_4	CCS
(0.2, 0.16, 0.16, 0.16, 0.16, 0.16)	0.049	0.054	0.051	0.048	0.05
(0.3, 0.14, 0.14, 0.14, 0.14, 0.14)	0.052	0.05	0.049	0.05	0.047
(0.4, 0.12, 0.12, 0.12, 0.12, 0.12)	0.051	0.05	0.049	0.048	0.039
(0.5, 0.10, 0.10, 0.10, 0.10, 0.10)	0.051	0.05	0.047	0.049	0.032
(0.6, 0.08, 0.08, 0.08, 0.08, 0.08)	0.052	0.054	0.049	0.048	0.027
(0.7, 0.06, 0.06, 0.06, 0.06, 0.06)	0.051	0.051	0.045	0.041	0.028
(0.8, 0.04, 0.04, 0.04, 0.04, 0.04)	0.049	0.048	0.044	0.036	0.028
(0.9, 0.02, 0.02, 0.02, 0.02, 0.02)	0.051	0.046	0.035	0.036	0.025

Table 6: Cell probabilities for 2 × 6 tables with ordered categories.

A.

	Effectiveness					
Drugs	Not effective	+	++	+++	++++	+++++
Placebo	0.8	0.04	0.04	0.04	0.04	0.04
Treatment	0.7	0.02	0.03	0.06	0.08	0.11

B.

	Effectiveness					
Drugs	Not effective	+	++	+++	++++	+++++
Placebo	0.8	0.04	0.04	0.04	0.04	0.04
Treatment	0.7	0.11	0.08	0.06	0.03	0.02

C.

	Effectiveness					
Drugs	Not effective	+	++	+++	++++	+++++
Placebo	0.8	0.04	0.04	0.04	0.04	0.04
Treatment	0.7	0.02	0.08	0.1	0.08	0.02

D.

	Effectiveness					
Drugs	Not effective	+	++	+++	++++	+++++
Placebo	0.8	0.04	0.04	0.04	0.04	0.04
Treatment	0.7	0.1	0.04	0.02	0.04	0.1

E.

	Effectiveness					
Drugs	Not effective	+	++	+++	++++	+++++
Placebo	0.8	0.04	0.04	0.04	0.04	0.04
Treatment	0.7	0.03	0.11	0.08	0.01	0.07

F.

	Effectiveness					
Drugs	Not effective	+	++	+++	++++	+++++
Placebo	0.8	0.04	0.04	0.04	0.04	0.04
Treatment	0.7	0.08	0.01	0.07	0.11	0.03

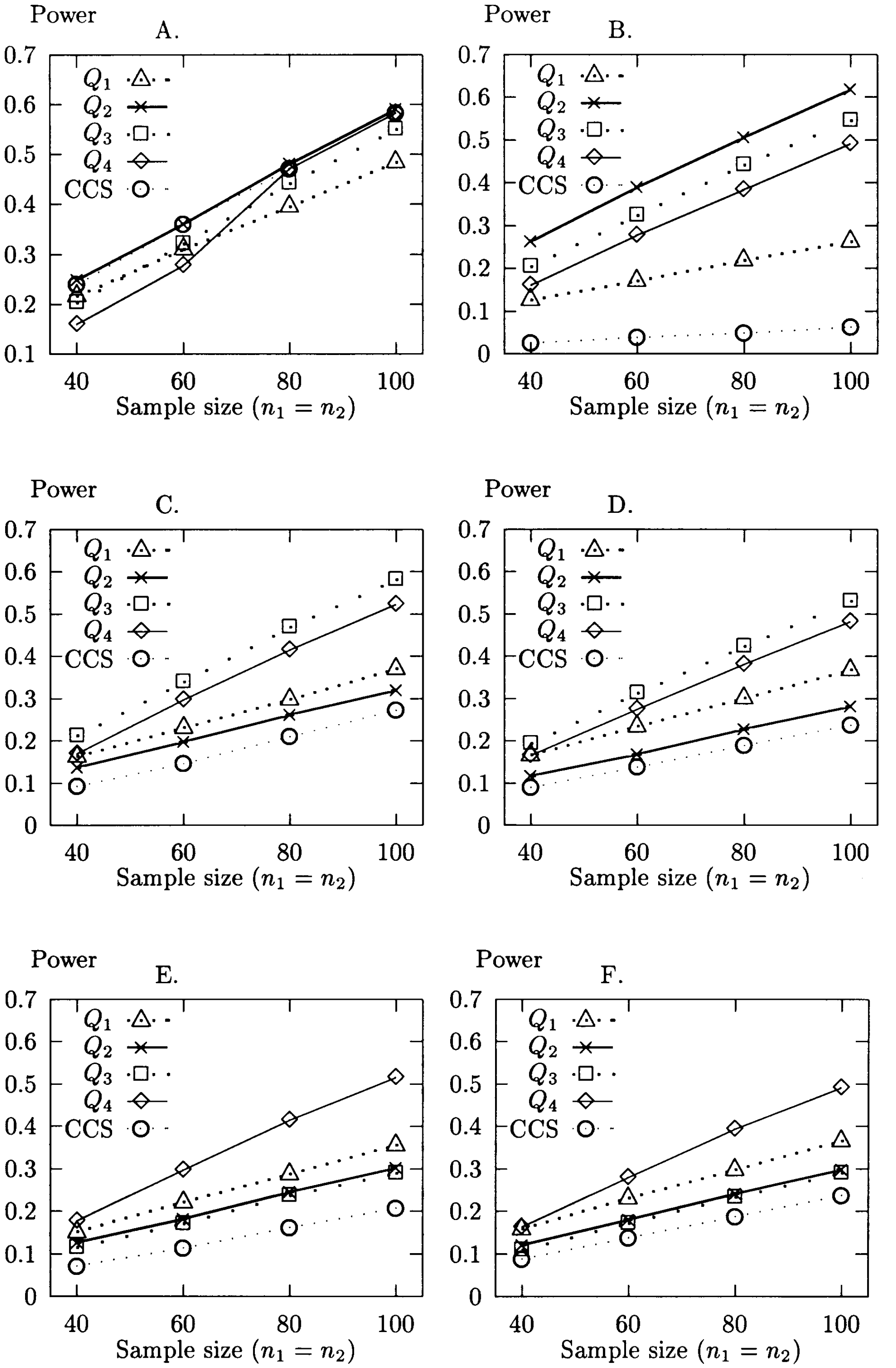

Figure 3: Power comparisons with the CCS test.

category increase and those of the other categories decrease the Type I errors of CCS test decrease considerably, whereas Type I error of the proposed test are close to 0.05.

To compare the powers of the proposed test with the CCS test, we conducted similar simulation using the cell probabilities of 2×6 tables given in Table 6. The distribution patterns of these cell probabilities are similar to Figure 2.

We assessed the power of the proposed test and the CCS test for the sample sizes $n_1 = n_2 = 40(20)100$. The powers obtained from Table 6A are plotted in Figure 3A; from Table 6B are in Figure 3B and so on. Figure 3A shows that the cumulative chi-square test has comparable powers with Q_2 test and that it has higher powers than all the other Q_t test. But the other figures show that the powers of the cumulative chi-square test are smaller than any Q_t test, $t = 1, 2, 3, 4$. This finding would be reasonable, since the empirical significance levels of the CCS test are substantially lower than the nominal test level (see Table 5). Furthermore, in general, the powers of the cumulative chi-square test are poor for such response that we selected. In particular, this test would be useless for these response pattern given in Table 6B, since for these cell probabilities, the differences of the observed and expected cumulative sum tends to be negligible for $k \geq 2$.

Also Figure 3 shows that for the responses given in Table 6A and 6B, the test with $t = 2$ provides the maximum powers; for the responses given in Table 6C and 6D, the test with $t = 3$ provides the maximum powers; and for the responses given in Tables 6E and 6F, the test with $t = 4$ provides the maximum powers.

5. Concluding Remarks

A class of statistics is constructed for testing ordered categorical data with non-linear responses in $2 \times k$ tables. The asymptotic distributions of the proposed statistic are obtained under the null and alternative hypotheses. The asymptotic powers of the test are compared, and also the exact powers of the tests and CCS test are examined.

Summarizing the results we may suggest the use of Q_2 test if the response pattern except for the first category is linear; of Q_3 test if the pattern of the response except for the first category is $\bigcup$ shaped or $\bigcap$ shaped; and Q_4 test if the pattern of the response is $\cap\cup$ shaped or $\cup\cap$ shaped.

Finally, we conclude that according to simulation, the proposed test is better than the CCS test for testing the non-linear responses in $2 \times k$ tables.

Acknowledgements

The authors are indebted to the Japanese Foundation for Multidisciplinary Treatment of Cancer. This study was also supported in part by Grant-in-Aid for Cooperative Research, Ministry of Education, Science and Culture, Japanese Government.

References

[1] JAYASEKARA, L., YANAGAWA, T. and TSUJITANI, M.: *A Location-Dispersion Test for $2 \times k$ Tables*, Bulletin of Informatics and Cybernetics, **26** (1994), 125-139.

[2] LEHMANN, E. L.: *NONPARAMETRICS: Statistical Methods Based on Ranks*, Holden Day, Inc., San Francisco, (1975).

[3] MANTEL, N. *Chi-Square tests with one degree of freedom: extensions of the Mantel-Haenszel procedure*, J. Amer. Statist. Assoc., **58** (1963), 690-700.

[4] NAIR, V. N.: *On Testing in Industrial Experiments with Ordered Categorical Data*, Technometrics, **28** (1986), 283-311.

[5] TAKEUCHI, K. and HIROTSU, C.: *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** (1982), 1-13.

[6] WILCOXON, F.: *Individual Comparisons by Ranking Methods*, Biometrics, **1** (1945), 80-83.

Received November 14, 1994
Revised January 10, 1995