

A TEST OF QUASI-INDEPENDENCE IN ORDINAL TRIANGULAR CONTINGENCY TABLES

Ming-Tien Tsai and Pranab K. Sen

Academia Sinica and University of North Carolina

Abstract. A new procedure for testing quasi-independence in an ordinal triangular contingency table is proposed as a generalization of Pearson's chi square test. The test is asymptotically equivalent in terms of power to the corresponding restricted likelihood ratio test for contiguous alternatives. Numerical results are also included in this study.

Key words and phrases: Chi square bar distribution, restricted likelihood ratio tests, positive likelihood ratio dependence.

1. Introduction

Many incomplete contingency tables contain cells which have structurally zero probabilities. For a good review, see Bishop, Fienberg and Holland (1975) where detailed references to the relevant literature have also been cited. The quasi-independence model, which is a generalization of the independence model, was formally introduced by Goodman (1968). It is most commonly used to analyze incomplete contingency tables (Bishop et al. (1975)). Goodness of fit tests of the quasi-independence model are usually based on the conventional Pearson chi square test (CST) statistic and likelihood ratio test (LRT) statistic (Goodman (1968), Bishop and Fienberg (1969)). Generally speaking, the two tests are asymptotically first order equivalent under a sequence of Pitman-type local alternatives. In practice, often, the categorical variables are ordinal. If the data represent changes that can only occur in one direction, then we have a triangular contingency table (see Mantel (1970) and others for examples). In general, there are four types of triangular tables: Upper-right (left) and lower-left (right) triangular tables. Because any of these types can be reduced to the other three by interchanging the column and row variables and/or by reversing the category orderings, we may without loss of generality consider only the upper-right triangular case. In order to motivate the proposed methods we consider the following two sets of data.

1. The relationship between small loops and whorls in finger-prints of the right

hand studied earlier by Waite (1915), Harris and Treloar (1927), Goodman (1968) and Sarkar (1989), presented in Table 1.

Table 1

Whorls	Small loops					
	5	4	3	2	1	0
0	45	179	211	204	144	78
1		32	80	126	153	106
2			15	55	92	130
3				7	38	125
4					26	104
5						50

2. The relationship between initial and final ratings on disability of stroke patients studied earlier by Bishop and Fienberg (1969), Mantel (1970) and Sarkar (1989), presented in Table 2.

Table 2

Initial state	Final state				
	E	D	C	B	A
E	8	15	12	23	11
D		1	4	10	9
C			4	4	6
B				5	4
A					5

The proposed method can also be applied to the ordinal block-triangular tables. For simplicity, we restrict our attention to ordinal triangular tables only. First we note that for classifications having ordinal random variables, the population values of the local log-odds ratios are either uniformly non-negative or uniformly non-positive (Goodman (1979), Patefield (1982)). Having more information about the alternatives, the conventional Pearson CST and LRT remain valid, but are not optimal or efficient any more (Bartholomew (1959)). Thus it is anticipated to construct other tests which have better power properties. Unfortunately, by utilizing the additional information about the alternatives, the corresponding restricted LRT for quasi-independence in an ordinal triangular table has a very complicated form. This motivates the construction of a unified large sample test which enjoys the same asymptotic optimal properties as the corresponding LRT but is computationally simpler. The proposed procedure also

remains valid and is asymptotically optimal for testing independence for an ordinal complete contingency table, and it will reduce to the conventional Pearson chi square testing procedure for the nominal one.

2. The Proposed Test

Let X and Y be two ordinal categorical random variables with respect to the common index set $\{1, \dots, I\}$, for some $I \geq 2$. Let $p_{ij} = P\{X = i, Y = j\}$, $i, j = 1, \dots, I$, and $\pi = ((p_{ij}))$. Then the local log-odds ratios are uniformly non-negative (Agresti (1984)) iff

$$p_{ij}p_{kl} \geq p_{il}p_{kj}, \quad \text{for all } i \leq k, j \leq l. \quad (2.1)$$

For an ordinal contingency table, owing to the ordering of the categories of X and Y , intuitively it is expected that there exists a strong ordinal relation between X and Y . Thus it is reasonable to assume that the population values of local log-odds ratios are either uniformly non-negative or uniformly non-positive for ordinal contingency tables (Goodman (1979), Patefield (1982)). For an upper-right triangular table, a model can not satisfy the reversal of (2.1) (Sarkar (1989)); thus, the population value of local log-odds ratio is expected to be uniformly non-negative [namely, (2.1) holds]. The interpretation of (2.1) obviously indicates a strong form of positive dependence between X and Y . (2.1) amounts to saying that the model $\pi = ((p_{ij}))$ is said to be positive dependent (via likelihood ratio) introduced by Lehmann (1966); it is also known as totally positive of order 2 (TP_2). TP_2 -dependence means that the conditional distributions of X given Y (or Y given X) have the monotone likelihood ratio property. For an upper-right triangular table, we may let $p_{ij} = 0$ iff $i > j$. If we assume that structurally non-zero probabilities are multiplicative (i.e. $p_{ij} = \alpha_i\beta_j$ for some positive parameters α_i and β_j), then $\pi = ((p_{ij}))$ is known in the literature as the quasi-independence model (Goodman (1968), Bishop and Fienberg (1969)). In a complete table the independence model minimizes ordinal association in the class of positive likelihood ratio dependence models with a fixed set of marginals, likewise the quasi-independence model minimizes ordinal association in the class of positive likelihood ratio dependence models with a fixed set of marginals for an ordinal upper-right triangular table (Sarkar (1989)). Hence for an ordinal upper-right triangular table, one is interested in testing the null hypothesis of quasi-independence against the alternative of positive likelihood ratio dependence, i.e.

$$H_0 : p_{ij}p_{kl} = p_{il}p_{kj}, \quad 1 \leq i \leq k \leq j \leq l \leq I, \quad (2.2)$$

versus

$$H_1 : p_{ij}p_{kl} \geq p_{il}p_{kj} \quad \text{with at least one strict inequality.} \quad (2.3)$$

Various tests for H_0 against H_1 under an ordinal complete contingency table have been considered in the literature [viz. Goodman (1979), Grove (1980), Hirotsu (1983), Shi (1991) and others]. The main purpose of this paper is to propose an asymptotically optimal test for this restricted alternative problem under the quasi-independence model. First, we make the following reparameterizations: Let

$$\begin{aligned} \boldsymbol{\theta} = (\boldsymbol{\eta}', \boldsymbol{\delta}', \boldsymbol{\lambda}')' \quad & \text{where } \boldsymbol{\eta} = (\eta_1, \dots, \eta_I)', \boldsymbol{\delta} = (\delta_1, \dots, \delta_I)' \\ & \text{and } \boldsymbol{\lambda} = (\lambda_{12}, \lambda_{13}, \dots, \lambda_{\overline{I-2} \overline{I-1}})' \end{aligned} \quad (2.4)$$

with

$$\begin{aligned} \eta_i = \ln p_{iI}, \quad i = 1, \dots, I, \quad \delta_j = \ln(p_{jj}/p_{jI}), \quad j = 1, \dots, I \quad \text{and} \\ \lambda_{ij} = \ln(p_{ij}p_{\overline{i+1}I}/p_{\overline{i+1}j}p_{iI}), \quad i = 1, \dots, I-2, \quad j = i+1, \dots, I-1. \end{aligned} \quad (2.5)$$

Also consider a $(I-1)(I-2)/2 \times (I^2+I+2)/2$ matrix \mathbf{A} defined by

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \vdots & \mathbf{0} & \vdots & \mathbf{A}_1 \end{bmatrix}, \quad (2.6)$$

where both the null matrices are of order $(I-1)(I-2)/2 \times I$ and \mathbf{A}_1 is a square matrix of order $(I-1)(I-2)/2$ with $\mathbf{A}_1 = (a_{ij})$ such that $a_{ii} = 1$, for $i = 1, \dots, (I-1)(I-2)/2$, $a_{\overline{i+1}i} = -1$, for $i = 1, \dots, (I-1)(I-2)/2$ but $i \neq I-2, 2I-5, \dots, I(I-3)/2$ and $a_{ij} = 0$ for other combinations of (i, j) , $1 \leq i, j \leq (I-1)(I-2)/2$. Then the null and the alternative hypotheses can be rewritten as

$$H_0 : \boldsymbol{\mu} = \mathbf{0} \quad (2.7)$$

versus

$$H_1 : \boldsymbol{\mu} \in \Gamma = \{\boldsymbol{\mu} \in R^{(I^2+I+2)/2}; \boldsymbol{\mu} \geq \mathbf{0}, \|\boldsymbol{\mu}\| > 0\}, \quad (2.8)$$

where $\boldsymbol{\mu} = \mathbf{A}\boldsymbol{\theta}$ and $\|\cdot\|$ denotes the Euclidean norm. With the above reformulation, we note that (2.7) relates to the vertex of the cone (the positive orthant space) and hence, under H_0 in (2.7), the parameter $\boldsymbol{\mu}$ lies on the boundary of the parameter space. Thus the usual asymptotic chi square distribution theory relating to the classical LRT (or the Pearson CST) procedure is not applicable. We need to take into account the restricted nature of the parameter space in the formulation of an appropriate test procedure. Consider an $I \times I$ upper-right triangular contingency table with frequencies $n_{ij} > 0$ iff $i \leq j$. Let the row and the column marginals be given by $n_{i\cdot} = \sum_{j=1}^I n_{ij}$, $1 \leq i \leq I$, and $n_{\cdot j} = \sum_{i=1}^I n_{ij}$, $1 \leq j \leq I$ respectively. Also, as is common in complete contingency tables, we assume that the n_{ij} have a multinomial distribution with cell

probabilities $p_{ij} > 0$ iff $1 \leq i \leq j \leq I$ and with sample size $n = \sum_{i=1}^I \sum_{j=i}^I n_{ij}$. Then the likelihood function with respect to $\theta = (\eta', \delta', \lambda')$, defined in (2.4)-(2.5), is

$$f(n_{ij}; 1 \leq i \leq j \leq I) = \left(n! / \prod_{1 \leq i \leq j \leq I} n_{ij}! \right) \exp \left\{ \sum_{i=1}^I n_i \eta_i + \sum_{j=1}^I n_{.j} \delta_j + \sum_{1 \leq i < j \leq I-1} \left(\sum_{k=1}^i n_{kj} \right) \lambda_{ij} \right\}. \tag{2.9}$$

Hence the log LRT statistic is of the form

$$\Lambda_n = \sup_{\mu \in \Gamma} \left[\sum_{i=1}^I n_i \eta_i + \sum_{j=1}^I n_{.j} \delta_j + \sum_{1 \leq i < j \leq I-1} \left(\sum_{k=1}^i n_{kj} \right) \lambda_{ij} \right] - \sup_{\mu=0} \left[\sum_{i=1}^I n_i \eta_i + \sum_{j=1}^I n_{.j} \delta_j + \sum_{1 \leq i < j \leq I-1} \left(\sum_{k=1}^i n_{kj} \right) \lambda_{ij} \right]. \tag{2.10}$$

An algebraic (closed) expression for Λ_n in terms of the sample quantities is generally difficult to obtain. We introduce a sequence $\{K_n\}$ of local alternatives under which we have tests of much simpler form. Let

$$K_n = \bigcup_{\gamma \in \Gamma^*} K_\gamma^{(n)} \text{ where } K_\gamma^{(n)} : \lambda = n^{-1/2} \gamma, \quad \gamma \in \Gamma^*, \tag{2.11}$$

and

$$\Gamma^* = \left\{ \gamma \in R^{(I-1)(I-2)/2}; \mathbf{A}_1 \gamma \geq \mathbf{0}, \|\mathbf{A}_1 \gamma\| \neq 0 \right\}. \tag{2.12}$$

Note that (2.11) relates to the usual Pitman-type (local) alternatives but confined to the restricted domain Γ^* . For finding an algebraic expression for (2.10) under $\{K_n\}$, let

$$\mathbf{S}_n = \mathbf{B} \mathbf{W}_n, \tag{2.13}$$

where \mathbf{B} is a lower triangular matrix of order $(I - 1)(I - 2)/2$ defined by

$$\mathbf{B} = \begin{bmatrix} \mathbf{I}_{I-2} & \mathbf{0} & \cdots & \mathbf{0} & 0 \\ \mathbf{0I}_{I-3} & \mathbf{I}_{I-3} & \cdots & \mathbf{0} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0I}_2 & \mathbf{0I}_2 & \cdots & \mathbf{I}_2 & 0 \\ 0 \cdots 01 & 0 \cdots 01 & \cdots & 01 & 1 \end{bmatrix}, \tag{2.14}$$

and

$$\mathbf{W}_n = n^{-1/2} (n_{ij} - n \hat{p}_{ij}), \quad 1 \leq i < j \leq I - 1, \tag{2.15}$$

where $\hat{\pi} = ((\hat{p}_{ij}))$, $1 \leq i < j \leq I-1$, is the maximum likelihood estimator (MLE) of π under the quasi-independence model. Various explicit forms of $n\hat{p}_{ij}$ were obtained by Goodman (1968), Bishop and Fienberg (1969) and Sarkar (1989). Note that under H_0 in (2.7) or $\{K_n\}$ in (2.11), $n^{-1/2}\{n_{ij} - n\hat{p}_{ij}\}$, $1 \leq i < j \leq I-1$, are all bounded in probability, so that by the Taylor expansion, (2.10) can be simplified to

$$\Lambda_n = \sup_{\gamma \in \Gamma^*} \left\{ \gamma' S_n + \frac{1}{2} \gamma' B \Sigma_n B' \gamma + o_p(1) \right\} \quad (2.16)$$

$$\Sigma_n = E_0 \{ \mathbf{W}_n \mathbf{W}_n' \} |_{\hat{\pi}} \quad (2.17)$$

with E_0 denoting the expectation under the null hypothesis. Even in this simplified form the solution in (2.16) may depend heavily on the form of Γ^* . Note that by virtue of (2.11), the central limit theorem (on the n_{ij}) and Le Cam's third lemma (Hájek and Šidák (1967)), we obtain that under $\{K_\gamma^{(n)}\}$ (for any fixed $\gamma \in \Gamma^*$)

$$S_n \xrightarrow{D} N_{(I-1)(I-2)/2}(B \Sigma B' \gamma, B \Sigma B'), \quad \Sigma = \lim_{n \rightarrow \infty} \Sigma_n. \quad (2.18)$$

Also for any given $\gamma \in \Gamma^*$, for testing H_0 versus $K_\gamma^{(n)}$, by the Neyman-Pearson lemma, an asymptotically most powerful test statistic is

$$T_n(\gamma) = \gamma' S_n / (\gamma' B \Sigma_n B' \gamma)^{1/2}, \quad \gamma \in \Gamma^*, \quad (2.19)$$

where the null hypothesis H_0 is rejected for large values of $T_n(\gamma)$. Moreover, by the definition (2.11), $\{K_n\}$ is the union of the component hypotheses $\{K_\gamma^{(n)}, \gamma \in \Gamma^*\}$, so the overall union-intersection (UI) test statistic for testing H_0 versus $\{K_n\}$ is given by

$$Q_n = \sup_{\gamma \in \Gamma^*} \left\{ \gamma' S_n / (\gamma' B \Sigma_n B' \gamma)^{1/2} \right\}. \quad (2.20)$$

In view of (2.16) and by imposing the condition $\gamma' B \Sigma_n B' \gamma = 1$, we thus have, under H_0 as well as $\{K_n\}$,

$$\Lambda_n = Q_n + \frac{1}{2} \quad \text{as } n \rightarrow \infty. \quad (2.21)$$

Thus we may conclude that Q_n shares the same asymptotically optimal properties as the restricted LRT Λ_n . Next, to obtain Q_n in (2.20), we need to maximize $\gamma' S_n$ subject to the equality $\gamma' B \Sigma_n B' \gamma = \text{constant}$ and the inequality constraint $A_1 \gamma > \mathbf{0}$. For this non-linear programming problem, the Kuhn-Tucker-Lagrange point formula (Hadley (1964)) yields the following result. Let

$$\mathbf{Z}_n = A_1 S_n \quad \text{and} \quad \mathbf{\Delta}_n = A_1 B \Sigma_n B' A_1', \quad (2.22)$$

and $K = \{1, 2, \dots, (I - 1)(I - 2)/2\}$. For any subset a of K , we denote its complementary subset by a' ; $\phi \subseteq a \subseteq K$, and also denote the cardinality of a by $|a|$. For each a , we partition \mathbf{Z}_n and $\mathbf{\Delta}_n$ as

$$\mathbf{Z}'_n = (\mathbf{Z}'_{n(a)}, \mathbf{Z}'_{n(a')}) \quad \text{and} \quad \mathbf{\Delta}_n = \begin{pmatrix} \mathbf{\Delta}_{n(aa)} & \mathbf{\Delta}_{n(aa')} \\ \mathbf{\Delta}_{n(a'a)} & \mathbf{\Delta}_{n(a'a')} \end{pmatrix}. \tag{2.23}$$

Let

$$\mathbf{Z}_{n(a:a')} = \mathbf{Z}_{n(a)} - \mathbf{\Delta}_{n(aa')} \mathbf{\Delta}_{n(a'a')}^{-1} \mathbf{Z}_{n(a')} \tag{2.24}$$

and

$$\mathbf{\Delta}_{n(aa:a')} = \mathbf{\Delta}_{n(aa)} - \mathbf{\Delta}_{n(aa')} \mathbf{\Delta}_{n(a'a')}^{-1} \mathbf{\Delta}_{n(a'a)}. \tag{2.25}$$

The test statistic Q_n in (2.20) is then given by

$$Q_n^2 = \sum_{\phi \subseteq a \subseteq K} \left\{ \mathbf{Z}'_{n(a:a')} \mathbf{\Delta}_{n(aa:a')}^{-1} \mathbf{Z}_{n(a:a')} \right\} 1(\mathbf{Z}_{n(a:a')} > \mathbf{o}) 1(\mathbf{\Delta}_{n(a'a')}^{-1} \mathbf{Z}_{n(a')} \leq \mathbf{o}), \tag{2.26}$$

where $1(B)$ stands for the indicator function of a set B . Note that although (2.26) is expressed as a sum over $2^{(I-1)(I-2)/2}$ possible terms, it is in fact a single term corresponding to the particular (random) a for which both the indicator functions are one. Further, by virtue of (2.18) under H_0 , \mathbf{S}_n is asymptotically $N_{(I-1)(I-2)/2}(\mathbf{0}, \mathbf{B}\mathbf{\Sigma}\mathbf{B}')$. Hence, proceeding as in Tsai (1993) [and omitting the details], we obtain that

$$\lim_{n \rightarrow \infty} P\{Q_n^2 \leq x | H_0\} = \sum_{r=0}^{(I-1)(I-2)/2} w_r P\{\chi_r^2 \leq x\} \quad \text{for every } x \in (0, \infty), \tag{2.27}$$

where χ_r^2 stands for a random variable having the central chi square distribution with r degrees of freedom (D.F.) ($\chi_0^2 = 0$); the non-negative weights $w_0, \dots, w_{(I-1)(I-2)/2}$ are defined by

$$w_r = \sum_{\{a:|a|=r\}} P\{\mathbf{Z}_{a:a'} > \mathbf{o}\} P\{\mathbf{\Delta}_{a'a}^{-1} \mathbf{Z}_{a'} \leq \mathbf{o}\}, \quad r=0, 1, \dots, (I - 1)(I - 2)/2; \tag{2.28}$$

and \mathbf{Z} has the $[(I - 1)(I - 2)/2]$ -variate normal distribution with zero mean and dispersion matrix $\mathbf{\Delta} = \mathbf{A}_1 \mathbf{B}\mathbf{\Sigma}\mathbf{B}' \mathbf{A}'_1$. In the literature, the right hand side of (2.27) is known as the chi square bar distribution (Robertson, Wright and Dykstra (1988)). Also as $\mathbf{\Delta}$ is a completely specified matrix; the weights $w_r, r = 1, \dots, (I - 1)(I - 2)/2$, may be computed (approximately) from a subroutine given by Bohrer and Chow (1978) for $r \leq 10$, which rests on a subroutine to estimate multivariate orthant probabilities as can be found in Sun (1988) for $r \leq$

9. Here, we develop a Fortran algorithm to estimate the w_r by incorporating the subroutine given by Evans and Schwartz (1986) which works well for evaluating a multivariate orthant probability when $r \geq 10$ (see Tables 5 and 6 in Section 3).

Modern resampling plans (especially, jackknifing and bootstrapping) (Wu (1986)) may be incorporated with advantage to provide suitable estimates of the w_r in (2.28). We may formulate this approach as follows.

(i) Recall that the $\Delta_{aa'}$ are all based on the model under the null hypothesis, and the marginal probability distributions dictate the probability law under this null hypothesis. The elements of $\Delta_{aa'}$ are all functions of the two sets of marginal probabilities, and hence, they can be estimated consistently from the two sample marginal counterparts.

(ii) These functions are nonlinear (typically, product-type), and hence, the plug-in estimators are biased, although the bias can be shown to be $O(n^{-1})$.

(iii) The classical jackknifing method can be used with advantage to reduce this bias to $O(n^{-2})$.

(iv) Use the jackknifed estimators of the $\Delta_{aa'}$ in the subsequent steps; these are denoted by $\hat{\Delta}_{aa'}^*$.

(v) Let $k = \binom{I-1}{2}$, and generate a large number (say, kM) of standard normal deviates, and group them into M sets of k vectors each. Denote these k -vectors by $\mathbf{X}_1^*, \dots, \mathbf{X}_M^*$ respectively.

(vi) Choose a matrix \mathbf{D}^* such that $\hat{\Delta}^* = \mathbf{D}^* \mathbf{D}^{*'}.$ Then let $\mathbf{Z}_i^* = \mathbf{D}^* \mathbf{X}_i^*$, for $i = 1, \dots, M$. Define a and the set K as in after (2.22), and for each \mathbf{Z}_i^* , define the partitions as in (2.23)-(2.25) with Δ replaced by $\hat{\Delta}^*$.

(vii) For each r ($0 \leq r \leq k$), count the frequencies for the events $[\mathbf{Z}_{a:a'(i)}^* \geq \mathbf{0}, \hat{\Delta}_{a'a'}^{*-1} \mathbf{Z}_{a'(i)}^* \leq \mathbf{0}]$, $a : |a| = r$, and these divided by M , yield the desired estimates of the w_r . Generation of the \mathbf{X}_i^* is inexpensive, and hence, a modern computer can take care of the estimation of the w_r by an adequate choice of M .

3. Discussions

Little is known about the exact power since the exact distribution of the proposed test is difficult to obtain for small and moderate sample sizes. As such, the asymptotic distribution theory is studied in this section and compared with existing methods. First we give some illustrations of the procedure proposed in this article. Following Goodman (1968), and Bishop and Fienberg (1969), the MLE of the unknown parameters np_{ij} , $1 \leq i \leq j \leq I$, under the quasi-independence model for the data in Tables 1 and 2 are given in Tables 3 and 4 respectively.

Table 3. Estimated frequencies for Table 1 under quasi-independence model

Whorls	Small loops					
	5	4	3	2	1	0
0	45.00	131.13	150.34	166.56	167.36	200.61
1		79.87	91.57	101.45	101.93	122.18
2			64.09	71.02	71.36	85.53
3				52.97	53.23	63.80
4					59.12	70.88
5						50.00

Table 4. Estimated frequencies for Table 2 under quasi-independence model

Initial state	Final state				
	E	D	C	B	A
E	8.00	11.48	11.94	21.92	15.66
D		4.52	4.69	8.63	6.16
C			3.37	6.20	4.43
B				5.25	3.75
A					5.00

In order to calculate the corresponding test statistic for each table, we first need to compute the corresponding covariance matrix Σ_n of \mathbf{W}_n (defined as in (2.17)). Although the MLE of np_{ij} , $1 \leq i \leq j \leq I$ can be explicitly found by an iterative formula (Goodman (1968)), we note that an explicit expression for Σ_n is quite difficult to obtain. To overcome this difficulty, we may replace the covariance matrix Σ_n with some other asymptotically equivalent ones which have simpler forms. Let

$$T_n = \sup_{\mathbf{A}_1 \gamma \neq 0} \left\{ \gamma' \mathbf{S}_n / (\gamma' \mathbf{B} \Sigma_n \mathbf{B}' \gamma)^{1/2} \right\}. \tag{3.4}$$

Then we have

$$T_n^2 = \mathbf{Z}'_n \mathbf{\Delta}_n^{-1} \mathbf{Z}_n = \mathbf{W}'_n \Sigma_n^{-1} \mathbf{W}_n, \tag{3.5}$$

which is the corresponding score test statistic for testing H_0 against global alternatives (i.e. $p_{ij}p_{kl} \neq p_{il}p_{kj}$, for all $i \leq k, j \leq l$). Let $n_{ij}^* = n^{-1/2}(n_{ij} - n\hat{p}_{ij})$, $1 \leq i \leq j \leq I$, and $\mathbf{n}^* = (n_{11}^*, \dots, n_{1I}^*, \dots, n_{II}^*)'$. By projecting \mathbf{n}^* onto a linear subspace of $R^{(I-1)(I-2)/2}$ such that all of the perturbations due to the equations $\sum_{i=1}^I n_{ij} = \sum_{i=1}^I n\hat{p}_{ij}$ and $\sum_{j=1}^I n_{ij} = \sum_{j=1}^I n\hat{p}_{ij}$, $1 \leq i \leq j \leq I$, are removed, we may rewrite the conventional Pearson CST statistic as

$$G_n^2 = \sum_{1 \leq i \leq j \leq I} (n_{ij} - n\hat{p}_{ij})^2 / (n\hat{p}_{ij}) = \mathbf{W}'_n (\Sigma_n^0)^{-1} \mathbf{W}_n, \tag{3.6}$$

where $(\Sigma_n^0)^{-1}$ is a positive definite matrix in probability and it can be explicitly expressed in terms of \hat{p}_{ij} for each I . For instance, if we write $\hat{p}_{ij} = m_{ij}^{-1}$, then $(\Sigma_n^0)^{-1}$ can be expressed as follows when $I = 6$

$$(\Sigma_n^0)^{-1} = \begin{bmatrix} a_1 & m_{16} & m_{16} & m_{16} & -m_{26} & -m_{26} & -m_{26} & 0 & 0 & 0 \\ m_{16} & a_2 & m_{16} & m_{16} & m_{33}+m_{36} & 0 & 0 & -m_{36} & -m_{36} & 0 \\ m_{16} & m_{16} & a_3 & m_{16} & 0 & m_{44}+m_{46} & 0 & m_{44}+m_{46} & 0 & -m_{46} \\ m_{16} & m_{16} & m_{16} & a_4 & 0 & 0 & m_{55}+m_{56} & 0 & m_{55}+m_{56} & m_{55}+m_{56} \\ -m_{26} & m_{33}+m_{36} & 0 & 0 & a_5 & m_{26} & m_{26} & -m_{36} & -m_{36} & 0 \\ -m_{26} & 0 & m_{44}+m_{46} & 0 & m_{26} & a_6 & m_{26} & m_{44}+m_{46} & 0 & -m_{46} \\ -m_{26} & 0 & 0 & m_{55}+m_{56} & m_{26} & m_{26} & a_7 & 0 & m_{55}+m_{56} & m_{55}+m_{56} \\ 0 & -m_{36} & m_{44}+m_{46} & 0 & -m_{36} & m_{44}+m_{46} & 0 & a_8 & m_{36} & -m_{46} \\ 0 & -m_{36} & 0 & m_{55}+m_{56} & -m_{36} & 0 & m_{55}+m_{56} & m_{36} & a_9 & m_{55}+m_{56} \\ 0 & 0 & -m_{46} & m_{55}+m_{56} & 0 & -m_{46} & m_{55}+m_{56} & -m_{46} & m_{55}+m_{56} & a_{10} \end{bmatrix}$$

where $a_1 = m_{12} + m_{16} + m_{22} + m_{26}, a_2 = m_{13} + m_{16} + m_{33} + m_{36}, a_3 = m_{14} + m_{16} + m_{44} + m_{46}, a_4 = m_{15} + m_{16} + m_{55} + m_{56}, a_5 = m_{23} + m_{26} + m_{33} + m_{36}, a_6 = m_{24} + m_{26} + m_{44} + m_{46}, a_7 = m_{25} + m_{26} + m_{55} + m_{56}, a_8 = m_{34} + m_{36} + m_{44} + m_{46}, a_9 = m_{35} + m_{36} + m_{55} + m_{56}, a_{10} = m_{45} + m_{46} + m_{55} + m_{56}$.

Since $\mathbf{W}_n \xrightarrow{D} N_{(I-1)(I-2)/2}(\mathbf{0}, \Sigma)$ as $n \rightarrow \infty$ under H_0 and $\lim_{n \rightarrow \infty} P\{G_n^2 \geq c | H_0\} = P\{\chi_{(I-1)(I-2)/2}^2 \geq c\}$ (Bishop and Fienberg (1969)), and following Ogasawara and Takahashi (1951) we have $\Sigma = \Sigma^0$, where $\Sigma^0 = \lim_{n \rightarrow \infty} \Sigma_n^0$. Thus, with the replacement of Σ_n by Σ_n^0 , we may conclude that the proposed test still has the same asymptotic optimal properties as the corresponding restricted LRT under a sequence of local alternatives $\{K_n\}$. Furthermore, from (2.20) and (3.4) we may note that Q_n^2 is a generalization of the score test statistic T_n^2 . Therefore with Σ_n replaced by Σ_n^0 our proposed test can be regarded as a generalization of the conventional Pearson CST.

With Σ_n replaced by Σ_n^0 , the corresponding test statistics, the non-negative weights $w_r, r = 1, \dots, (I - 1)(I - 2)/2$ and asymptotic p -values for Tables 1 and 2 are calculated and presented in Tables 5 and 6 respectively. From the result of Table 5, not only Goodman’s conclusion that the null hypothesis of quasi-independence model should be rejected can be supported, but also we may conclude that there exists a negative likelihood ratio dependence for the underlying model in the finger-print data set. For the data set of Table 2, based on the results of conventional Pearson CST and LRT, Bishop and Fienberg supported the quasi-independence model. However, Table 6 suggests that there may exist a positive likelihood ratio dependence for the same underlying model. In passing, we note that $Q_n^2 \leq T_n^2$. Moreover, if we let $\alpha = \lim_{n \rightarrow \infty} P\{T_n^2 \geq \tau_\alpha | H_0\} =$

$P\{\chi_{(I-1)(I-2)/2}^2 \geq \tau_\alpha\} = \lim_{n \rightarrow \infty} P\{Q_n^2 \geq \rho_\alpha | H_0\} = \sum_{r=0}^{(I-1)(I-2)/2} w_r P\{\chi_r^2 \geq \rho_\alpha\}$, then by the property of the log concavity of χ_j^2 in $j, \alpha \in (0, 1)$ (Das Gupta and Sarkar (1984)), we have $\tau_\alpha > \rho_\alpha, \alpha \in (0, 1)$. However, little information can be extracted from this fact as far as a power comparison of our proposed test and the conventional Pearson CST is concerned. Fortunately, the admissibility of our proposed tests follows from a theorem of Eaton (1970). The theorem implies that the class of tests for H_0 versus H_1 (defined in (2.7)-(2.8)) which have closed convex acceptance regions containing the dual cone $\mathcal{C} = \{\mathbf{v} \in R^{(I^2+I+2)/2}; \mathbf{v}' \Sigma_n^{-1} \boldsymbol{\mu} \leq \mathbf{0}, \boldsymbol{\mu} \geq \mathbf{0}\}$ of Γ , constitutes an essentially complete class. However, unlike our proposed test, the conventional Pearson CST is inadmissible. Thus for the problem of testing the quasi-independence model in ordinal triangular tables, the proposed test is preferred.

Table 5. The test statistics, weights and asymptotic p -values of the proposed test for Table 1.

r	weights w_r	$Q_n^2 = 387.42$ $P\{\chi_r^2 \geq Q_n^2\}$
0	0.003580	0
1	0.040920	0
2	0.138286	0
3	0.277543	0
4	0.284959	0
5	0.168430	0
6	0.062173	0
7	0.012991	0
8	0.001690	0
9	0.000100	0
10	0.000003	0
	$\sum_{r=0}^{10} w_r$ $= 0.990675$	$p\text{-value} \doteq 0$

Table 6. The test statistic, weights and asymptotic p -value of the proposed test for Table 2.

r	weights w_r	$Q_n^2 = 7.2291$ $P\{\chi_r^2 \geq Q_n^2\}$
0	0.029767	0
1	0.176077	0.007173
2	0.358623	0.026929
3	0.307001	0.064943
4	0.110957	0.124264
5	0.017304	0.204150
6	0.000867	0.300018
	$\sum_{r=0}^6 w_r =$ 1.000592	$p\text{-value}$ $= 0.048438$

Sarkar (1989) exploited the ordinal nature of row and column variables to narrow the alternative hypothesis and proposed a linear test in terms of the difference of sample proportions of concordances and discordances. His test can be regarded as a generalized version of Kendall's tau test. In terms of the notation in Section 2, note that Sarkar's test is equivalent to the LRT of testing the null hypothesis $H_{0\alpha} : \mathbf{1}'\boldsymbol{\eta} = 0$ versus $H_{1\alpha} : \mathbf{1}'\boldsymbol{\eta} > 0$, where $\boldsymbol{\eta} = \mathbf{A}_1\boldsymbol{\lambda}$. Hence

Sarkar's test is admissible and asymptotically consistent. On the other hand, some other admissible and asymptotically consistent tests against an alternative that corresponds to some specific models may also be considered, such as testing for quasi-independence against Goodman's quasi-uniform association model (Goodman (1979)). Obviously, this problem is equivalent to the problem of testing the null hypothesis $H_0^* : \boldsymbol{\eta} = \mathbf{0}$ versus $H_1^* : \boldsymbol{\eta} = \eta \mathbf{1}, \eta > 0$. For such a testing problem the corresponding LRT statistic, being a linear function of \mathbf{W}_n , is the most suitable one for this particular one-sided simple alternative according to the Neyman-Pearson lemma. Here we remark that the advantage of such a class of linear test statistics (see (2.19)) is that it has an asymptotically normal distribution, so that the asymptotic critical level can be computed easily. The linear tests considered above are all sensitive to the restricted alternatives. They are most powerful against the corresponding specific alternative but they may perform poorly for most of alternatives of Γ . In most circumstances, it is unlikely that one knows what the specific alternatives are, so it is probably more sensible to use a test which has reasonable power properties for all alternatives. If we still restrict ourselves to a class of linear tests, we can do better by using a class of asymptotically most stringent somewhere most powerful tests (Abelson and Tukey (1963), Schaafsma and Smid (1966)). However, numerical studies show that the asymptotic power of the optimal linear test is still rather poor near the boundary of the alternative parameter space Γ (Abelson and Tukey (1963)). On the other hand our proposed test (asymptotically equivalent to the restricted LRT) performs robustly over the entire parameter space of restricted alternatives (Tsai, Sen and Yang (1994)).

Acknowledgement

Thanks are due to the chair editor, an associate editor and the referees for their helpful comments of the manuscript.

References

- Abelson, R. P. and Tukey, J. W. (1963). Efficient utilization of non-numerical information in quantitative analysis: General theory and the case of simple order. *Ann. Math. Statist.* **34**, 1347-1369.
- Agresti, A. (1984). *Analysis of Ordinal Categorical Data*. John Wiley, New York.
- Bartholomew, D. J. (1959). A test of homogeneity for ordered alternatives I, II. *Biometrika* **46**, 36-48, 328-335.
- Bishop, Y. M. M. and Fienberg, S. E. (1969). Incomplete two-dimensional contingency tables. *Biometrics* **25**, 119-128.
- Bishop, Y. M. M, Fienberg, S. E. and Holland, P. W. (1975). *Discrete Multivariate Analysis*. MIT Press, Cambridge, MA.

- Bohrer, R. and Chow, W. (1978). Algorithm AS 122. Weights for one-sided multivariate inference. *Appl. Statist.* **27**, 100-104.
- Das Gupta, S. and Sarkar, S. K. (1984). On TP_2 and log-concavity. *Inequalities in Statistics and Probability*, Vol.5 (Edited by Y. L. Tong (1982)), 54-58, The Institute of Mathematical Statistics, Hayward, CA.
- Eaton, M. L. (1970). A complete class theorem for multidimensional one-sided alternatives. *Ann. Math. Statist.* **41**, 1884-1888.
- Evans, M. and Schwartz, T. (1986). Monte Carlo computation of some multivariate normal probabilities. Technical Report No.7, University of Toronto, Canada.
- Goodman, L. A. (1968). The analysis of cross-classified data: Independence, quasi-Independence, and interactions in contingency tables with or without missing entries. *J. Amer. Statist. Assoc.* **63**, 1091-1131.
- Goodman, L. A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *J. Amer. Statist. Assoc.* **74**, 537-552.
- Grove, D. M. (1980). A test of independence against a class of ordered alternatives in a $2 \times C$ contingency table. *J. Amer. Statist. Assoc.* **75**, 454-459.
- Hadley, G. (1964). *Non-linear and Dynamic Programming*. Addison-Wesley, Reading, MA.
- Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests*. Academic Press, New York.
- Harris, J. A. and Treloar, A. E. (1927). On a limitation in the applicability of the contingency coefficient. *J. Amer. Statist. Assoc.* **22**, 460-472.
- Hirotsu, C. (1983). Defining the pattern of association in two-way contingency tables. *Biometrika* **70**, 579-589.
- Lehmann, E. H. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37**, 1137-1153.
- Mantel, N. (1970). Incomplete contingency tables. *Biometrics* **26**, 291-304.
- Ogasawara, T. and Takahashi, M. (1951). Independence of quadratic forms in normal system. *J. Sci. Hiroshima University* **15**, 1-9.
- Patefield, W. M. (1982). Exact tests for trends in ordered contingency tables. *Appl. Statist.* **31**, 32-43.
- Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*. John Wiley, New York.
- Sarkar, S. K. (1989). Quasi-independence in ordinal triangular contingency tables. *J. Amer. Statist. Assoc.* **84**, 592-597.
- Schaafsma, W. and Smid, L. J. (1966). Most stringent somewhere most powerful tests against alternatives restricted by a number of linear inequalities. *Ann. Math. Statist.* **37**, 1161-1172.
- Shi, N. Z. (1991). A test of homogeneity of odds ratios against order restrictions. *J. Amer. Statist. Assoc.* **86**, 154-158.
- Sun, H. J. (1988). A Fortran subroutine for computing normal orthant probabilities of dimensions up to nine. *Comm. in Statist., Part B – Simulation and Computation* **17**, 1097-1111.
- Tsai, M.-T. (1993). UI score tests for some restricted alternatives in exponential families. *J. Multivariate Anal.* **45**, 305-323.
- Tsai, M.-T., Sen, P. K. and Yang, Y. H. (1994). Power-robustness of likelihood ratio (union-intersection score) tests for some restricted alternative problems. *Statistics & Decisions* **12**, 231-244.
- Waite, H. (1915). Association of finger-prints. *Biometrika* **10**, 421-478.

Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Ann. Statist.* **14**, 1261-1295.

Institute of Statistical Science, Academia Sinica, Taipei 11529, Taiwan.

Departments of Biostatistics and Statistics, University of North Carolina, Chapel Hill, NC 27599-7400, U.S.A.

(Received September 1993; accepted November 1994)