Journal of the Korean
Statistical Society
Vol. 25, No. 3, 1996

Inference for Order Restrictions on Odds in $2 \times k$ Contingency Tables

Myongsik Oh [1]

Abstract

In the analysis of contingency table with ordered categories, the relationship between odds for adjacent categories has received considerable interest. We consider likelihood ratio tests of independence against an order restriction on odds in $2 \times k$ contingency tables.

Key Words : Chi-bar-square distribution; Contingency table; Isotonic regression; Likelihood ratio test; Local odds ratio.

1. INTRODUCTION

In contingency tables, the relationship between odds for adjacent categories has been of considerable interest. Throughout this paper we assume ordered contingency table. One of the most frequently encountered relationships is positive (negative) likelihood ratio dependence. Such dependence concepts were brought to prominance by Lehmann (1966). This is also called total positivity of order 2 (TP_2). Sociologists often call it trend. In contingency tables this restriction can be expressed in terms of local odds ratios.

[1]Department of Statistics, Pusan University of Foreign Studies, Pusan 608-738, Korea.

Let p_{ij} be the cell probability of ith row and jth column for $i = 1, 2,\ j = 1, 2, \ldots, k$, and $\sum_{ij} p_{ij} = 1$. Local odds ratios are defined as

$$\theta_j = \frac{p_{1j}\, p_{2,j+1}}{p_{1,j+1}\, p_{2,j}},\ j = 1, 2, \ldots, k-1. \tag{1.1}$$

The independence hypothesis, H_0, can be expressed in terms of θ_j's by

$$H_0:\ \theta_1 = \theta_2 = \cdots = \theta_{k-1} = 1. \tag{1.2}$$

If $\theta_j \geq (\leq) 1, j = 1, 2, \ldots, k-1$, or equivalently p_{1j}/p_{2j} is decreasing (increasing) in j, a contingency table is positively (negatively) likelihood ratio dependent. Grove (1984) considered the likelihood ratio test for testing H_0 against the hypothesis of positive likelihood ratio dependence.

In some situations one might be interested in testing the hypothesis which imposes different type of restriction on local odds ratios. For example, $\theta_j \geq 1$ for $j = 1, 2, \ldots, a$, $\theta_j \leq 1$ for $j = a+1, \ldots, k-1$. For a more general setting let J_1 and J_2 be the two mutually disjoint subsets of $J = \{1, 2, \ldots, k-1\}$ such that $J_1 \bigcup J_2 = J$. We consider the hypothesis

$$H_1:\ \theta_j \leq 1 \text{ if } j \in J_1,\ \theta_j \geq 1 \text{ if } j \in J_2. \tag{1.3}$$

If J_1 is empty and $J_2 = J$ then the restriction is equivalent to positive likelihood ratio dependence.

In this paper we develop likelihood ratio tests for H_0 against H_1 in $2 \times k$ contingency tables with some different types of sampling schemes. These sampling schemes are namely, multinomial, product multinomial, binomial and Poisson models. We, however, focus mainly on multinomial model.

We state these four sampling schemes briefly. Suppose that each cell count has an independent Poisson distribution. This is Poisson model. If we fix the total count, each cell no longer has Poisson distribution. We will call this multinomial model. If we fix row (column) total, each row (column) has multinomial distribution. We will call this product multinomial model. Since the column product multinomial model in a $2 \times k$ contingency tables is just a product of k independent binomial models, we will call this binomial model.

In section 2, we discuss the maximum likelihood estimation of cell probabilities under H_1. In section 3, we derive the asymptotic distribution of the test statistics. In section 4, we will consider inferences under other sampling schemes. Finally in section 5, a real data set is analyzed to illustrate the test procedure developed in section 3.

2. MAXIMUM LIKELIHOOD ESTIMATION

Suppose $0 < p_{ij} < 1, i = 1,2, j = 1,2,\ldots,k$, and $\sum_{ij} p_{ij} = 1$. Let n be the sample size and $\hat{p}_{ij}$'s be relative frequencies. It is convenient to use a one-to-one transformation of the parameter space by setting

$$\psi_j = \frac{p_{1j}}{p_{1j}+p_{2j}}, \quad \phi_j = p_{1j}+p_{2j}, \text{ for } j = 1,2,\ldots,k.$$

Then $p_{1j} = \phi_j\psi_j$, $p_{2j} = \phi_j(1-\psi_j)$, $j = 1,2,\ldots,k$ and the local odds ratios, θ_j, are expressed by

$$\theta_j = \frac{\psi_j\,(1-\psi_{j+1})}{(1-\psi_j)\,\psi_{j+1}}, \quad j = 1,2,\ldots,k-1.$$

It is easy to show that the basic restrictions are

$$\text{(A)} \qquad 0 \le \psi_j \le 1, \ 0 \le \phi_j \le 1, \ j = 1,2,\ldots,k,$$

$$\text{(B)} \qquad \sum_{j=1}^{k} \phi_j = 1.$$

and the likelihood function becomes

$$\left[\prod_{j=1}^{k} \psi_j^{n\hat{p}_{1j}}(1-\psi_j)^{n\hat{p}_{2j}}\right] \cdot \left[\prod_{j=1}^{k} \phi_j^{n(\hat{p}_{1j}+\hat{p}_{2j})}\right]. \tag{2.1}$$

The first part in (2.1), which involves only ψ_j's, can be viewed as a product of k independent binomial likelihood functions while the second part is a multinomial likelihood function. The unconstrained maximum likelihood estimator's (MLE's) of ψ_j and ϕ_j are given by $\hat{\psi}_j = \hat{p}_{1j}/(\hat{p}_{1j}+\hat{p}_{2j})$ and $\hat{\phi}_j = \hat{p}_{1j}+\hat{p}_{2j}$ provided $\hat{p}_{1j}+\hat{p}_{2j} > 0$.

The independence hypothesis (1.2) is equivalent to

$$H_0 : \ \psi_1 = \psi_2 = \cdots = \psi_k. \tag{2.2}$$

Then the MLE's of ψ_j's and ϕ_j's under H_0 are given by ψ_j° and ϕ_j°, where

$$\psi_1^\circ = \psi_2^\circ = \cdots = \psi_k^\circ = \sum_{j=1}^{k} \hat{p}_{1j} \text{ and } \phi_j^\circ = \hat{p}_{1j}+\hat{p}_{2j} \text{ for } j = 1,2,\ldots,k,$$

and thus $p_{ij}^\circ = (\hat{p}_{1j}+\hat{p}_{2j})\sum_{\ell=1}^{k} \hat{p}_{i\ell}$ for $i = 1,2, j = 1,2,\ldots,k$.

We next consider the MLE's under H_1, which is expressed in terms of ψ_j's and ϕ_j's as

$$H_1 : \ \psi_j \le \psi_{j+1} \text{ if } j \in J_1 \text{ and } \psi_j \ge \psi_{j+1} \text{ if } j \in J_2. \tag{2.3}$$

Consider a binary relation $\preceq$ on $X = \{1, 2, \ldots, k\}$ such that

$$j \preceq j+1 \text{ if } j \in J_1 \text{ and } j+1 \preceq j \text{ if } j \in J_2.$$

It is not difficult to show that the binary relation $\preceq$ is a partial order on X. Let $\mathsf{A} = (\psi_1, \psi_2, \ldots, \psi_k)$. Then restriction (2.3) is equivalent to

$$H_1 : \ \mathsf{A} \text{ is isotonic with respect to } \preceq . \tag{2.4}$$

Now the estimation problem is to find ψ_j's and ϕ_j's which maximize (2.1) subject to (2.4) together with basic restrictions (A) and (B). Since (2.4) does not involve ϕ_j's we can maximize (2.1) by maximizing the two parts separately. As in Example 1.5.1 of Robertson, Wright and Dykstra (1988) the first part is a problem encountered frequently in bioassy. The second part is just a multinomial problem.

Let $\mathcal{I} = \{\mathbf{x} \in \mathbf{R}^k : x_j \le x_{j+1} \text{ if } j \in J_1 \text{ and } x_j \ge x_{j+1} \text{ if } j \in J_2\}$. We note that $\mathcal{I}$ is a closed convex cone in $\mathbf{R}^k$. It follows from Theorem 1.5.2 of Robertson *et al.* (1988) that the restricted MLE, A^*, of A is the isotonic regression of $\hat{\mathsf{A}}$ with weight $\hat{\mathbf{p}}_1 + \hat{\mathbf{p}}_2 = (\hat{p}_{11} + \hat{p}_{21}, \hat{p}_{12} + \hat{p}_{22}, \ldots, \hat{p}_{1k} + \hat{p}_{2k})$ onto the convex cone $\mathcal{I}$. We express this fact by writing $\mathsf{A}^* = E_{\hat{\mathbf{p}}_1 + \hat{\mathbf{p}}_2}(\hat{\mathsf{A}}|\mathcal{I})$. The Minimum Lower Sets Algorithm (MLSA, see pp 24-25 of Robertson *et al.* 1988) can be used to compute $E(\cdot|\mathcal{I})$. On the other hand, the MLE, ϕ_j^*, of ϕ_j is $\hat{p}_{1j} + \hat{p}_{2j}$. Then we have the following theorem.

Theorem 1. If $\hat{p}_{1j} + \hat{p}_{2j} > 0, \ j = 1, 2, \ldots, k$, then the MLE of $\mathbf{p}$ under H_1 is given by $\mathbf{p}^*$, where

$$\begin{aligned} p_{1j}^* &= (\hat{p}_{1j} + \hat{p}_{2j}) E_{\hat{\mathbf{p}}_1 + \hat{\mathbf{p}}_2} \left(\frac{\hat{\mathbf{p}}_1}{\hat{\mathbf{p}}_1 + \hat{\mathbf{p}}_2} | \mathcal{I} \right)_j , \\ p_{2j}^* &= (\hat{p}_{1j} + \hat{p}_{2j}) E_{\hat{\mathbf{p}}_1 + \hat{\mathbf{p}}_2} \left(\frac{\hat{\mathbf{p}}_2}{\hat{\mathbf{p}}_1 + \hat{\mathbf{p}}_2} | \mathcal{A} \right)_j , \end{aligned}$$

for $j = 1, 2, \ldots, k$, $\hat{\mathbf{p}}_1 = (\hat{p}_{11}, \hat{p}_{12}, \ldots, \hat{p}_{1k})$, $\hat{\mathbf{p}}_2 = (\hat{p}_{21}, \hat{p}_{22}, \ldots, \hat{p}_{2k})$, $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_k + y_k)$, $\mathbf{x}/\mathbf{y} = (x_1/y_1, \ldots, x_k/y_k)$, $\mathcal{A} = \{\mathbf{x} \in \mathbf{R}^k : -\mathbf{x} \in \mathcal{I}\}$, and $E(\cdot)_j$ denotes the jth component of $E(\cdot)$.

By noting that the projection operator $E_{\mathbf{w}}(\mathbf{x}|\cdot)$ is continuous in both $\mathbf{w}$ and $\mathbf{x}$ (see Corollary to Theorem 1.4.4 of Robertson *et al.* 1988) and the strong law of large numbers, we have the strong consistency of the estimator. Specifically, if H_1 is true, then $Pr[\lim_{n\to\infty} \mathbf{p}^* = \mathbf{p}] = 1$.

3. THE LIKELIHOOD RATIO TEST

We consider the likelihood ratio statistic for testing the null hypothesis H_0 against the alternative H_1. The likelihood ratio statistic is

$$\begin{aligned}\Lambda &= \frac{\sup_{\mathbf{p}\in H_0} L(\mathbf{p})}{\sup_{\mathbf{p}\in H_1} L(\mathbf{p})} \\ &= \frac{\prod_{j=1}^k (\psi_j^\circ)^{n\hat{p}_{1j}} (1-\psi_j^\circ)^{n(1-\hat{p}_{1j})} \prod_{j=1}^k (\phi_j^\circ)^{n(\hat{p}_{1j}+\hat{p}_{2j})}}{\prod_{j=1}^k (\psi_j^*)^{n\hat{p}_{1j}} (1-\psi_j^*)^{n(1-\hat{p}_{1j})} \prod_{j=1}^k (\phi_j^*)^{n(\hat{p}_{1j}+\hat{p}_{2j})}} \\ &= \prod_{j=1}^k \left(\frac{\psi_j^\circ}{\psi_j^*}\right)^{n\hat{p}_{1j}} \left(\frac{1-\psi_j^\circ}{1-\psi_j^*}\right)^{n\hat{p}_{2j}}\end{aligned}$$

since $\phi_j^\circ = \phi_j^*$. Let $T = -2\ln\Lambda$. The test rejects H_0 for large value of

$$T = 2n\sum_{j=1}^k \{\hat{p}_{1j}\ln\psi_j^* + \hat{p}_{2j}\ln(1-\psi_j^*) - \hat{p}_{1j}\ln\psi_j^\circ - \hat{p}_{2j}\ln(1-\psi_j^\circ)\}.$$

We find the asymptotic distribution of T under H_0. By expanding $\ln\psi_j^*$ and $\ln\psi_j^\circ$ about $\hat{\psi}_j$, and $\ln(1-\psi_j^*)$ and $\ln(1-\psi_j^\circ)$ about $1-\hat{\psi}_j$ and using properties of isotonic regressions (see Theorem 1.3.3 of Robertson *et al.* 1988) we have

$$T = \sum_{j=1}^k \left\{(\hat{\psi}_j - \psi_j^\circ)^2 \left(\frac{n\hat{p}_{1j}}{\alpha_j^2} + \frac{n\hat{p}_{2j}}{\gamma_j^2}\right) - (\hat{\psi}_j - \psi_j^*)^2 \left(\frac{n\hat{p}_{1j}}{\beta_j^2} + \frac{n\hat{p}_{2j}}{\delta_j^2}\right)\right\}$$

where $\alpha_j(\beta_j)$ is between $\hat{\psi}_j$ and ψ_j° ($\hat{\psi}_j$ and ψ_j^*) and $\gamma_j(\delta_j)$ is between $1-\hat{\psi}_j$ and $1-\psi_j^\circ$ ($1-\hat{\psi}_j$ and $1-\psi_j^*$). When H_0 is true $\alpha_j,\ \beta_j \to \psi_j$ and $\gamma_j,\ \delta_j \to 1-\psi_j$, with probability one, since $\psi_j^\circ,\ \psi_j^* \to \psi_j$. Also note that $(p_{1j}+p_{2j})^2/(p_{1j}p_{2j}) = 1/(p_{1\cdot}p_{2\cdot})$ for $j = 1,2,\ldots,k$ under H_0, where $p_{i\cdot} = \sum_{j=1}^k p_{ij},\ i = 1,2$. Hence T is approximately equal to

$$\begin{aligned}&\sum_{j=1}^k \left\{(\hat{\psi}_j - \psi_j^\circ)^2 - (\hat{\psi}_j - \psi_j^*)^2\right\} \left\{\frac{n\hat{p}_{1j}}{\hat{\psi}_j^2} + \frac{n\hat{p}_{2j}}{(1-\hat{\psi}_j)^2}\right\} \\ &\approx \frac{n}{p_{1\cdot}p_{2\cdot}} \sum_{j=1}^k \left\{(\hat{\psi}_j - \psi_j^\circ)^2 - (\hat{\psi}_j - \psi_j^*)^2\right\} (\hat{p}_{1j} + \hat{p}_{2j}). \qquad (3.1)\end{aligned}$$

It follows from Theorem 1.3.2 of Robertson *et al.* (1988) that (??) is equal to

$$\frac{n}{p_{1\cdot}p_{2\cdot}}\sum_{j=1}^{k}(\psi_j^* - \psi_j^\circ)^2(\hat{p}_{1j}+\hat{p}_{2j})$$
$$= \frac{n}{p_{1\cdot}p_{2\cdot}}\sum_{j=1}^{k}\left[E_{\hat{\mathbf{p}}_1+\hat{\mathbf{p}}_2}(\hat{\mathbf{A}}|\mathcal{C})_j - E_{\hat{\mathbf{p}}_1+\hat{\mathbf{p}}_2}(\hat{\mathbf{A}}|\mathcal{I})_j\right]^2(\hat{p}_{1j}+\hat{p}_{2j})$$
$$= \frac{1}{p_{1\cdot}p_{2\cdot}}\sum_{j=1}^{k}\left[E_{\hat{\mathbf{p}}_1+\hat{\mathbf{p}}_2}(\sqrt{n}(\hat{\mathbf{A}}-\psi\cdot\mathbf{1})|\mathcal{C})_j - E_{\hat{\mathbf{p}}_1+\hat{\mathbf{p}}_2}(\sqrt{n}(\hat{\mathbf{A}}-\psi\cdot\mathbf{1})|\mathcal{I})_j\right]^2(\hat{p}_{1j}+\hat{p}_{2j}),$$

where $\psi = \psi_1 = \cdots = \psi_k$, $\hat{\mathbf{A}} = (\hat{\psi}_1, \hat{\psi}_2, \ldots, \hat{\psi}_k)$, $\mathbf{1} = (1, 1, \ldots, 1)$ and $\mathcal{C} = \{\mathbf{x} \in \mathbf{R}^k : x_1 = x_2 = \cdots = x_k\}$.

By the straightforward use of delta method we show that $\sqrt{n}(\hat{\mathbf{A}} - \psi\cdot\mathbf{1})$ converges in distribution to $N(\mathbf{0}, \Sigma_{\mathbf{p}})$, where $\Sigma_{\mathbf{p}} = [\sigma_{ij}]_{k\times k}$ with

$$\sigma_{ij} = \begin{cases} \frac{p_{1j}p_{2j}}{(p_{1j}+p_{2j})^3} & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Thus T converges in distribution to

$$\frac{1}{p_{1\cdot}p_{2\cdot}}\sum_{j=1}^{k}\left[E_{\hat{\mathbf{p}}_1+\hat{\mathbf{p}}_2}(\mathbf{Y}|\mathcal{C})_j - E_{\hat{\mathbf{p}}_1+\hat{\mathbf{p}}_2}(\mathbf{Y}|\mathcal{I})_j\right]^2(p_{1j}+p_{2j}),$$

where $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_k)$, Y_j is independent normal variate with mean 0 and variance $(p_{1j}+p_{2j})^3/(p_{1j}p_{2j})$, and $\hat{\mathbf{p}}_1 + \hat{\mathbf{p}}_2 = (\hat{p}_{11}+\hat{p}_{21}, \ldots, \hat{p}_{1k}+\hat{p}_{2k})$. We note that $(p_{1j}p_{2j})/(p_{1j}+p_{2j})^3 = (p_{1\cdot}p_{2\cdot})/(p_{1j}+p_{2j})$ for $j = 1, 2, \ldots, k$ under H_0. Noting that $E_{\mathbf{w}}(\mathbf{x}|\cdot) = E_{a\cdot\mathbf{w}}(\mathbf{x}|\cdot)$ for $a > 0$ and appealing to Theorem 2.3.1 of Robertson *et al.* (1988) lead to the following.

Theorem 2. If H_0 is true and $n \to \infty$, then for all t

$$\lim_{n\to\infty} Pr[T \geq t] = \sum_{\ell=1}^{k} P(\ell, k; \mathbf{p}_1 + \mathbf{p}_2, \preceq)Pr[\chi^2_{\ell-1} \geq t], \tag{3.2}$$

where χ^2_ν denotes an independent chi-square random variable with ν degrees of freedom ($\chi^2_0 \equiv 0$) and $P(\ell, k; \mathbf{p}_1 + \mathbf{p}_2, \preceq)$ is the probability that $E_{\mathbf{p}_1+\mathbf{p}_2}(\mathbf{Y}|\mathcal{I})$ takes on ℓ distinct values.

We note that H_0 does not specify the value of $p_{1j} + p_{2j}$. The asymptotic null distribution of T does depend upon unknown parameter $\mathbf{p}_1 + \mathbf{p}_2$ through

$P(\ell, k; \mathbf{p}_1 + \mathbf{p}_2, \preceq)$, which we call level probabilities. We recommend that the quantity (3.2) be approximated by

$$\sum_{\ell=1}^{k} P(\ell, k; \mathbf{p}_1^\circ + \mathbf{p}_2^\circ, \preceq) Pr[\chi^2_{\ell-1} \geq t], \tag{3.3}$$

where $\mathbf{p}_1^\circ$ and $\mathbf{p}_2^\circ$ are restricted MLE under H_0. Oh (1994) studied methods for approximating the null hypothesis distributions of several test statistics by using an estimate of the unknown quantity on which the null distribution depends. It is shown that this method generally provides a very good approximation to the true asymptotic null distribution of test statistic.

For general partial order it is not easy to compute exact level probabilities $P(\ell, k; \mathbf{w}, \preceq)$. Robertson *et al.* (1988), however, provide the table of level probabilities for the simple order and a simple tree ordering when the weights $\mathbf{w}$ are equal. Shi(1988) and Singh and Wright (1993) gave level probabilities for a unimodal (umbrella) order and a simple loop ordering, respectively. For the unequal weight cases equal weight approximation is found to be adequate for most practical purposes. The approximation techniques for well known orderings have been developed by Robertson and Wright (1983), Wright and Tran (1985), Lucas, Robertson and Wright (1989) and Singh and Wright (1993). Although it has no practical use we can compute equal weight level probabilities for other than well known orderings. We may implement a computer program by using the recursive relationship in Theorem 2.4.1 of Robertson *et al.* (1988). An equal weight approximation technique can be employed to compute unequal weight level probabilities.

4. OTHER SAMPLING SCHEMES

We begin this section with discussing estimation procedure under H_1 with three different types of sampling schemes. Oh(1995) studied estimation procedures of cell probabilities in $2 \times$ k contingency table under negative dependence restrictions with various sampling schemes. The three estimation problems discussed in this section can be solved by slight modification of some estimation procedure develpoed in Oh(1995).

First consider the product multinomial model. Suppose the first row is distributed as $(m, p_1, p_2, \ldots, p_k)$ and the second row is distributed as $(n, q_1, q_2, \ldots, q_k)$. Then the local odds ratios are defined as $\theta_j = (p_j\, q_{j+1})/(p_{j+1}\, q_j)$, $j = 1, 2, \ldots, k-1$. It is not difficult to show that if $\theta_j = 1$ for $j = 1, \ldots, k-1$ then

$\mathbf{p} = \mathbf{q}$. Dykstra, Kochar and Robertson (1995) studied the likelihood ratio statistic for testing H_0 against H_1, where $J_1 = \{1, 2, \ldots, k-1\}$ and $J_2 = \emptyset$. Following their reparametrization scheme ($\psi_j = mp_j/(mp_j + nq_j)$, $\phi_j = mp_j + nq_j$ for $j = 1, 2, \ldots, k$), the restricted MLE's of $\mathbf{p}$ and $\mathbf{q}$ are given by $\mathbf{p}^*$ and $\mathbf{q}^*$, where

$$\begin{aligned} \mathbf{p}^* &= \frac{m\hat{\mathbf{p}} + n\hat{\mathbf{q}}}{m} E_{m\hat{\mathbf{p}}+n\hat{\mathbf{q}}}\left(\frac{m\hat{\mathbf{p}}}{m\hat{\mathbf{p}} + n\hat{\mathbf{q}}}|\mathcal{I}\right), \\ \mathbf{q}^* &= \frac{m\hat{\mathbf{p}} + n\hat{\mathbf{q}}}{n} E_{m\hat{\mathbf{p}}+n\hat{\mathbf{q}}}\left(\frac{n\hat{\mathbf{q}}}{m\hat{\mathbf{p}} + n\hat{\mathbf{q}}}|\mathcal{A}\right), \end{aligned}$$

$\mathcal{I}$ and $\mathcal{A}$ are defined in Theorem 2.1.

We next consider the binomial model. Suppose each column has a binomial distribution with sample size n_j and success probability p_j. Then the null hypothesis is $H_0 : p_1 = p_2 = \cdots = p_k$ and the alternative hypothesis becomes

$$H_1 : \mathbf{p} \text{ is isotonic with respect to } \preceq,$$

where $\preceq$ is defined in section 2. The restricted MLE of $\mathbf{p}$ under H_1 is given by $\mathbf{p}^*$, where $\mathbf{p}^* = E_{\mathbf{n}}(\hat{\mathbf{p}}|\mathcal{I})$, and $\mathbf{n} = (n_1, n_2, \ldots, n_k)$.

We next consider Poisson model. Suppose $X_{ij\ell}$, $\ell = 1, 2, \ldots, n$, be random sample from Poisson population with mean λ_{ij}. It is also convenient to use a one-to-one transformation of the parameter space by setting

$$\psi_j = \frac{\lambda_{1j}}{\lambda_{1j} + \lambda_{2j}}, \ \phi_j = \lambda_{1j} + \lambda_{2j}, \ , j = 1, 2, \ldots, k.$$

Then $\lambda_{1j} = \psi_j\phi_j$ and $\lambda_{2j} = \phi_j(1-\psi_j)$. Using the similar estimation procedure as discussed in section 2 we can show that the restricted MLE's of λ_{ij}'s are given by λ_{ij}^*'s, where

$$\begin{aligned} \lambda_{1j}^* &= \frac{\sum_{\ell=1}^{n}(x_{1j\ell} + x_{2j\ell})}{n} E_{\mathbf{x}_1+\mathbf{x}_2}(\hat{\mathsf{A}}|\mathcal{I})_j, \text{ and} \\ \lambda_{2j}^* &= \frac{\sum_{\ell=1}^{n}(x_{1j\ell} + x_{2j\ell})}{n} E_{\mathbf{x}_1+\mathbf{x}_2}(\mathbf{1} - \hat{\mathsf{A}}|\mathcal{A})_j, \end{aligned}$$

$\hat{\mathsf{A}} = ((\sum_{\ell=1}^{n} x_{11\ell})/(\sum_{\ell=1}^{n}(x_{11\ell}+x_{21\ell})), \ldots, (\sum_{\ell=1}^{n} x_{1k\ell})/(\sum_{\ell=1}^{n}(x_{1k\ell}+x_{2k\ell})))$ and $\mathbf{x}_1 + \mathbf{x}_2 = (\sum_{\ell=1}^{n}(x_{11\ell} + x_{21\ell}), \ldots, \sum_{\ell=1}^{n}(x_{1k\ell} + x_{1k\ell}))$. We note that we can relax the assumption of the equal sample sizes. For the unequal sample case we use the following reparametrization; $\psi_j = n_{1j}\lambda_{1j}/(n_{1j}\lambda_{1j} + n_{2j}\lambda_{2j})$,

$\phi_j = n_{1j}\lambda_{1j} + n_{2j}\lambda_{2j}$, $j = 1, 2, \ldots, k$. The estimation and distribution theory is similar to the case of the equal sample sizes.

Next we consider the asymptotic null distribution of the test statistics for each sampling scheme. Let T be the test statistics. For each sampling scheme T can be defined appropriatley as in multinomial sampling case. Then we can show that the asymptotic null distribution of T is

$$\lim Pr[T \geq t] = \sum_{\ell=1}^{k} P(\ell, k; \mathbf{w}, \preceq) Pr[\chi^2_{\ell-1} \geq t].$$

where

$$\mathbf{w} = \begin{cases} \mathbf{p}(=\mathbf{q}) & \text{for product multinomial sampling,} \\ (\lim_{n\to\infty} \frac{n_1}{n}, \ldots, \lim_{n\to\infty} \frac{n_k}{n}) & \text{for binomial sampling,} \\ (\lambda_{11} + \lambda_{21}, \cdots, \lambda_{1k} + \lambda_{2k}), & \text{for Poisson sampling.} \end{cases}$$

It is of interest to observe that the asymptotic null distribution of T for binomial sampling scheme does not depend on parameter $\mathbf{p}$ but on column sample sizes (ratios to total count) while others depend upon parameter not on sample size.

5. AN EXAMPLE

To illustrate the testing procedure discussed in earlier sections we examine a data set discussed in McCullagh (1980).

Table 1. Quality of right eye vision in men and women

	vision quality				
	highest 1	high 2	low 3	lowest 4	total
men	1053	782	893	514	3242
women	1976	2256	2456	789	7477
total	3029	3038	3349	1303	10719

From the table 1 we see that women are relatively more concentrated in the middle categories while men have higher proportion in two extreme categories. We frequently encounter this type of dependency in practice. We may set the following hypothesis which is more relaxed than the aforementioned hypothesis. For higher vision quality, gender and vision quality are positively dependent, and for low vision quality they are negatively dependent. We note that gender is not an ordered variable. If the two rows are interchanged the hypothesis should be changed accordingly.

In terms of local odds ratios the hypothesis is expressed by

$$H_1 : \ \theta_1 \geq 1, \ \theta_3 \leq 1.$$

The corresponding partial order, $\preceq$, on $X = \{1,2,3,4\}$ is $2 \preceq 1$ and $3 \preceq 4$.

Table 2 shows the computational details. The test statistic is 129.362. To find p-value we need to know the level probabilities. Since the value of $\mathbf{p}_1 + \mathbf{p}_2$ is not specified in H_0 we use (3.3). The level probabilities for this data set are, however, obtained easily. Since the hypothesis does not impose any restriction between $\{1,2\}$ and $\{3,4\}$, we first find level probabilities on $\{1,2\}$ and $\{3,4\}$ separately and then combine them. We know that $P(1,2;\mathbf{w}) = P(2,2;\mathbf{w}) = 1/2$ for any positive weight $\mathbf{w}$ of dimension 2. See page 77 of Robertson *et al.* (1988). Combining level probabilities we have $P(1,4;\mathbf{w}) = 0$, $P(2,4;\mathbf{w}) = 1/4$, $P(3,4;\mathbf{w}) = 1/2$, and $P(4,4;\mathbf{w}) = 1/4$. Now the p-value is obtained as

$$\begin{aligned} 0.000 &\approx 0 \cdot Pr[\chi_0^2 \geq 129.36] + 1/4 \cdot Pr[\chi_1^2 \geq 129.36] \\ &+ 1/2 \cdot Pr[\chi_2^2 \geq 129.36] + 1/4 \cdot Pr[\chi_3^2 \geq 129.36]. \end{aligned}$$

The test result provides that the H_1 is true.

Table 2. Computations of MLE's

$\hat{p}_{1j}$	$\hat{p}_{2j}$	$\hat{\psi}_j$	$\hat{\phi}_j$	ψ_j^*	p_{1j}°	p_{2j}°	ψ_j°
.0982	.1843	.3476	.2826	.3476	.0855	.1971	.3025
.0730	.2105	.2575	.2834	.2575	.0857	.1977	.3024
.0833	.2291	.2666	.3124	.2666	.0945	.2179	.3025
.0480	.0736	.3947	.1216	.3947	.0368	.0848	.3026

$\hat{\phi}_j = \phi_j^* = \phi_j^\circ$

REFERENCES

(1) Dykstra, R. L., S. C. Kochar, and T. Robertson (1995). Inference for likelihood ratio ordering in the two-sample problem, *Journal of the American Statistical Association*, **90**, 1034-1040.

(2) Grove, D. M. (1984). Positive association in a two-way contingency table: Likelihood ratio tests, *Communications in Statistics, Part A-Theory and Methods*, **13**, 931-945.

(3) Lehmann. E. L. (1966). Some concepts of dependence, *Annals of Mathematical Statistics*, **37**, 1137-1153.

(4) Lucas, L. A., T. Robertson, and F. T. Wright (1989). Approximating the level probabilities for a unimodal ordering, *Communications in Statistics, Part B-Simulation and Computations*, **18**, 1401-1420.

(5) McCullagh, P. (1980). Regression models for ordinal data, *Journal of the Royal Statistical Society, Series B*, **42**, 109-142.

(6) Oh, M. (1994). Statistical tests concerning a set of multinomial parameters under order restrictions: Approximations to null hypotheses distributions, Ph. D. Thesis, The University of Iowa, Iowa City, Iowa.

(7) Oh, M. (1995). On maximum likelihood estimation of cell probabilities in $2 \times k$ contingency tables under negative dependence restrictions with various sampling schemes, *Communications in Statistics, Part A-Theory and Methods*, **24**(8), 2127-2143.

(8) Robertson, T., F. T. Wright, and R. L. Dykstra (1988). *Order Restricted Statistical Inference*, Wiley, Chichester.

(9) Robertson, T., and F. T. Wright (1983). On approximation of the level probabilities and associated distributions in order restricted inference, *Biometrika*,**70**, 597-606.

(10) Singh, B., and F. T. Wright (1993). The level probabilities for a simple loop ordering, *Annals of Institute of Statistical Mathematics*, **45**, 279-292.

(11) Wright, F. T., and T. A. Tran (1985). Approximating the level probabilities in order restricted inference: the simple tree ordering, *Biometrika*, **72**, 429-439.