Journal of Korean
Data & Information Science Society
2003, Vol. 14, No.2 pp. 413~428

Statistical Inference Concerning Local Dependence between Two Multinomial Populations

Myongsik Oh[1)]

Abstract

If a restriction is imposed only to a (proper) subset of parameters of interest, we call it a local restriction. Statistical inference under a local restriction in multinomial setting is studied. The maximum likelihood estimation under a local restriction and likelihood ratio tests for and against a local restriction are discussed. A real data is analyzed for illustrative purpose.

Keywords : Chi-bar-square, local dependence, likelihood ratio ordering, stochastic ordering, uniform stochastic ordering.

1. Introduction

In statistical analysis of $2\times k$ contingency table, the most frequently used statistical concept is dependence. This dependence concept is closely related to various types of stochastic ordering, such as usual stochastic ordering, uniform stochastic ordering (hazard rate ordering for continuous random variables), likelihood ratio ordering and dominated ordering, among others. Douglas, Fienberg, Lee, Sampson, and Whitaker (1990) provides an excellent illustration of these relationships. Statistical inference under such ordering have been studied by many researchers. Robertson and Wright (1981) studied statistical inference concerning stochastic ordering between two multinomial populations, Dykstra, Kochar, and Robertson (1991, 1995) provided statistical inferential procedure for uniform stochastic ordering and likelihood ratio ordering. Park, Lee, and Robertson (1998) studied likelihood ratio test for uniform stochastic ordering in two or more discrete distribution with the same support. Chang (1993) studied dominance ordering for

1) Associate Professor, Department of Statistics, Pusan University of Foreign Studies, Busan, 608-730, Korea
E-mail : moh@pufs.ac.kr

two multinomial parameters.

We consider the following example which has motivated of this study. In the academic years of 1997 and 1998, 280 and 208 students were admitted to the College of Oriental Studies, Pusan University of Foreign Studies respectively. For each academic year stdents were classified into 10 groups according to their high school ranks. There are originally 15 scales in high school ranks but only top 10 scales are used. The smaller the number, the higher the school rank. It is well known that the academic performance of first year college students is closely related to their high school ranks. The higher the school rank, the better GPA in the first year. A school official claims that for high school rank 5 or above the portion of admitted students up to certain high school rank has increased in 1998 compared to 1997. Let p_i the proportions of ith group for admitted students in 1998 and q_i be the proportions of ith group for admitted students in 1997. Then the school official's assertion may be hypothesized as a local stochastic ordering, i.e., $\sum_{j=1}^{i} p_j \ge \sum_{j=1}^{i} q_j$ for $i=1,\cdots,5$. We are interested in testing $H_0: p_i = q_i$ for $i=1,\cdots,10$ versus $H_1: \sum_{j=1}^{i} p_j \ge \sum_{j=1}^{i} q_j$ for $i=1,\cdots,5$.

We call this type of restriction local dependence. As we discussed before there are several types of dependence. First suppose local odds ratios, $p_i q_{i+1}/p_{i+1} q_i$, are greater than or equal to 1 for $i=1,\cdots,k_0$. This is equivalent to that p_i/q_i is nondecreasing in $i=1,\cdots,k_0$, which shows likelihood ratio ordering locally. Second, suppose $p_i \sum_{\ell=i+1}^{k} q_\ell / q_i \sum_{\ell=i+1}^{k} p_\ell \ge 1$ for $i=1,\cdots,k_0$. It is equivalent to that $(1-\sum_{\ell=1}^{i} p_\ell)/(1-\sum_{\ell=1}^{i} q_\ell)$ is nonincreasing in $i=1,\cdots,k_0$, which shows uniform stochastic ordering locally. Third, suppose $\sum_{\ell=1}^{i} p_\ell \sum_{\ell=i+1}^{k} q_\ell / \sum_{\ell=1}^{i} q_\ell \sum_{\ell=i+1}^{k} p_\ell \ge 1$ for $i=1,\cdots,k_0$. This is equivalent to $\sum_{\ell=1}^{i} p_\ell \ge \sum_{\ell=1}^{i} q_\ell$ for $i=1,\cdots,k_0$, which is similar to stochastic ordering locally. Finally, let $p_i \ge q_i$ for $i=1,\cdots,k_0$, which is similar to dominated ordering locally. It is appropriate to use the term ``local'' in front of each ordering since the transformations of parameter preserve the corresponding ordering as we will see later.

In this paper we are going to discuss the statistical inferences when one of four types of local dependence is imposed to a $2\times k$ contingency table. In Section 2, maximum likelihood (ML) estimate of cell probabilities under the local dependence restriction is discussed. Transformation of parameter space is required to find cell probabilities. In section 3, likelihood ratio tests (LRT's) of equality of two parameter against local dependence is studied. The asymptotic null distributions of test statistics are derived.

2. Estimation

In this section we are going to investigate the estimation procedure for four types of local dependence restrictions. For each type of restriction, the key step is reparametrization of parameter space before applying the existing estimation procedure. However the reparametrization scheme for each restriction is different from ordering to ordering.

Let $\mathbf{p}=(p_1,p_2,\cdots,p_k)$ and $\mathbf{q}=(q_1,q_2,\cdots,q_k)$ be the probability vectors. Suppose that $\widehat{\mathbf{p}}$ and $\widehat{\mathbf{q}}$ are the relative frequencies of successes corresponding to independent random samples of sizes m and n from the $\widehat{\mathbf{p}}$ and $\widehat{\mathbf{q}}$ populations, respectively. Let N=m+n and let $\overline{\mathbf{p}}$ and $\overline{\mathbf{q}}$ be the ML estimates of $\mathbf{p}$ and $\mathbf{q}$, respectively. Note that we use the same notation for restricted ML estimate regardless of restriction type unless it causes confusion.

2.1 Local stochastic ordering

Suppose $\sum_{j=1}^{i} p_j \ge \sum_{j=1}^{i} q_j$ for $i=1,\cdots,k_0$. Let $a_i=p_i$, $b_i=q_i$ for $i=1,\cdots,k_0$, $a_{k_0+1}=\sum_{j=k_0+1}^{k} p_j$, $b_{k_0+1}=\sum_{j=k_0+1}^{k} q_j$, $\phi_i=p_i/a_{k_0+1}$, $\tau_i=q_i/b_{k_0+1}$ for $i=k_0+1,\cdots,k$. Then the basic restriction becomes $\sum_{i=1}^{k_0+1} a_i=1$, $\sum_{i=1}^{k_0+1} b_i=1$, $\sum_{i=k_0+1}^{k}\phi_i=1$, $\sum_{i=k_0+1}^{k}\tau_i=1$ and $0<a_i,b_i,\phi_i,\tau_i<1$. The local stochastic ordering restriction becomes

$$\sum_{j=1}^{i} a_j \ge \sum_{j=1}^{i} b_j \text{ for } i=1,\cdots,k_0, \text{ and } \sum_{j=1}^{k_0+1} a_j = \sum_{j=1}^{k_0+1} b_j. \tag{2.1}$$

The likelihood function becomes

$$\left[\prod_{i=1}^{k_0} a_i^{m\widehat{p}_i} \cdot a_{k_0+1}^{\sum_{j=k_0+1}^{k} m\widehat{p}_j} \cdot \prod_{i=1}^{k_0} b_i^{n\widehat{q}_i} \cdot b_{k_0+1}^{\sum_{j=k_0+1}^{k} n\widehat{q}_j}\right] \cdot \left[\prod_{i=k_0+1}^{k} \phi_j^{m\widehat{p}_j}\tau_j^{n\widehat{q}_j}\right]. \tag{2.2}$$

We observe that no restrictions except (2.1) relate a_i, b_i, ϕ_i and τ_i to each other. The maximum of (2.2) under (2.1) is obtained by maximizing the two parts (each in bracket) separately.

First consider the latter part in (2.2). Since no restriction other than basic restrictions are imposed each is just a usual multinomial model and hence $\bar{\phi}_i = \hat{\phi}_i = m_i / \sum_{j=k_0+1}^{k} m_j$ and $\bar{\tau}_i = \hat{\tau}_i = n_i / \sum_{j=k_0+1}^{k} n_j$, for $i = k_0+1, \cdots, k$. For the first part we can use estimation procedure in Robertson and Wright (1981). The restricted are given by

$$\bar{a} = \hat{a} E_{\hat{a}} \left(\frac{m\hat{a} + n\hat{b}}{N\hat{a}} | A \right),$$
$$\bar{b} = \hat{b} E_{\hat{b}} \left(\frac{m\hat{a} + n\hat{b}}{N\hat{a}} | I \right) = -\hat{b} E_{\hat{b}} \left(-\frac{m\hat{a} + n\hat{b}}{N\hat{a}} | A \right),$$

where $A = \{ x \in R^{k_0+1} : x_1 \geq x_2 \geq \cdots \geq x_{k_0+1} \}$, $I = \{ x \in R^{k_0+1} : -x \in A \}$, and all vector operations are componentwise. Evaluation of p and q at $a = \bar{a}$, $b = \bar{b}$, $\phi_i = \bar{\phi}_i$, $\tau_i = \bar{\tau}_i$ for $i = k_0+1, \cdots, k$ gives the restricted ML estimates $\bar{p}$ and $\bar{q}$.

Suppose $p = q$. This is equivalent to $a_i = b_i$ for $i = 1, \cdots, k_0+1$ and $\phi_i = \tau_i$ for $i = k_0+1, \cdots, k$. Then the ML estimates under this restriction are given by

$$\tilde{a}_i = \tilde{b}_i = \frac{m\hat{p}_i + n\hat{q}_i}{N}, \text{ for } i = i = 1, \cdots, k_0,$$
$$\tilde{a}_{k_0+1} = \tilde{b}_{k_0+1} = \frac{m \sum_{j=k_0+1}^{k} \hat{p}_i + n \sum_{j=k_0+1}^{k} \hat{q}_i}{N},$$
$$\tilde{\phi}_i = \tilde{\tau}_i = \frac{m\hat{p}_j + n\hat{q}_j}{\sum_{j=k_0+1}^{k} (m\hat{p}_j + n\hat{q}_j)}, \text{ for } i = k_0+1, \cdots, k.$$

2.2 Local uniform stochastic ordering

Suppose $(1 - \sum_{\ell=1}^{i} p_\ell)/(1 - \sum_{\ell=1}^{i} q_\ell)$ is nonincreasing in $i = 1, \cdots, k_0$. For this restriction we don't need another transformation of parameter space other than the transformation used for the case $k_0 = k$. Let $\theta_{1j} = \sum_{\ell=j+1}^{k} p_\ell / \sum_{\ell=j}^{k} p_\ell$ and $\theta_{2j} = \sum_{\ell=j+1}^{k} q_\ell / \sum_{\ell=j}^{k} q_\ell$ for $j = 1, \cdots, k-1$. Then $p_1 = 1 - \theta_{11}$, $p_j = (1 - \theta_{1j}) \prod_{\ell=1}^{j-1} \theta_{1\ell}$ for $j = 2, \cdots, k-1$, $p_k = \prod_{\ell=1}^{k-1} \theta_{1\ell}$, $q_1 = 1 - \theta_{21}$, $q_j = (1 - \theta_{2j}) \prod_{\ell=1}^{j-1} \theta_{2\ell}$ for $j = 2, \cdots, k-1$, $q_k = \prod_{\ell=1}^{k-1} \theta_{2\ell}$. Since

$\sum_{\ell=j}^{k} p_\ell = \prod_{\ell=1}^{j-1} \theta_{1\ell}$, $\sum_{\ell=j}^{k} q_\ell = \prod_{\ell=1}^{j-1} \theta_{2\ell}$ for $j=2,\cdots,k$, the restriction becomes

$$\theta_{1j} \le \theta_{2j} \text{ for } j=1,\cdots,k_0-1 \qquad (2.3)$$

and the likelihood function is

$$\prod_{j=1}^{k-1}\left[\theta_{1j}^{\sum_{\ell=j+1}^{k} m\hat{p}_\ell}(1-\theta_{1j})^{m\hat{p}_j}\cdot\theta_{2j}^{\sum_{\ell=j+1}^{k} n\hat{q}_\ell}(1-\theta_{2j})^{n\hat{q}_j}\right]. \qquad (2.4)$$

Note that for $j=k_0,\cdots,k-1$ no restriction is imposed between θ_{1j} and θ_{2j}. The basic restriction is $0<\theta_{ij}<1$ for $i=1,2,\ j=1,\cdots,k-1$. Since the restriction does not relate θ_{ij} for different j, the maximum of (2.4) under (2.3) can be achieved by maximizing j-1 terms (each in bracket) separately. Then restricted ML estimates are

$$\begin{aligned}
&\bar{\theta}_{1j}=\hat{\theta}_{1j},\ \bar{\theta}_{2j}=\hat{\theta}_{2j} && \text{if } \hat{\theta}_{1j}\le\hat{\theta}_{2j}, && \text{for } j=1,\cdots,k_0-1,\\
&\bar{\theta}_{1j}=\bar{\theta}_{2j}=\tilde{\theta}_{1j}=\tilde{\theta}_{2j} && \text{if } \hat{\theta}_{1j}>\hat{\theta}_{2j}, && \text{for } j=1,\cdots,k_0-1,\\
&\bar{\theta}_{1j}=\hat{\theta}_{1j},\ \bar{\theta}_{2j}=\hat{\theta}_{2j}, && && \text{for } j=k_0,\cdots,k-1,
\end{aligned}$$

where for $j=1,\cdots,k-1$,

$$\tilde{\theta}_{1j}=\tilde{\theta}_{2j}=\frac{m\sum_{\ell=j+1}^{k}\hat{p}_\ell+n\sum_{\ell=j+1}^{k}\hat{q}_\ell}{m\sum_{\ell=j}^{k}\hat{p}_\ell+n\sum_{\ell=j}^{k}\hat{q}_\ell}.$$

Evaluation of p and q at $\theta_{ij}=\bar{\theta}_{ij}$ for $i=1,2$ and $j=1,\cdots,k-1$ gives the restricted ML estimates $\bar{p}$ and $\bar{q}$. The ML estimate of θ_{ij} under restriction $p=q$ is $\tilde{\theta}_{ij}$.

2.3 Local likelihood ratio ordering

Suppose p_i/q_i is nondecreasing in $i=1,\cdots,k_0$. Let $a_{k_0}=\sum_{i=1}^{k_0} p_i$, $b_{k_0}=\sum_{i=1}^{k_0} q_i$ and $a_i=p_i, b_i=q_i$ for $i=k_0+1,\cdots,k$. Let

$$\theta_i=\frac{m' p_i/a_{k_0}}{m' p_i/a_{k_0}+n' q_i/b_{k_0}},\ \phi_i=m' p_i/a_{k_0}+n' q_i/b_{k_0} \text{ for } i=1,\cdots,k_0,$$

where $m'=m\sum_{i=1}^{k_0}\hat{p}_i$ and $n'=n\sum_{i=1}^{k_0}\hat{q}_i$. The basic restrictions become

(1) $0<a_i, b_i<1$ for $i=k_0,\cdots,k$, (2) $\sum_{i=k_0}^{k} a_i=\sum_{i=k_0}^{k} b_i=1$,

(3) $0\le\theta_i\le1$, $\phi_i\ge0$ for $i=1,\cdots,k_0$, (4) $\sum_{i=1}^{k_0}\phi_i=m'+n'$, and (5)

$\sum_{i=1}^{k_0} \theta_i \phi_i = m'$.

The local likelihood ratio ordering is equivalent to

$$\theta_1 \le \cdots \le \theta_{k_0}. \tag{2.5}$$

The likelihood function is proportional to

$$\left[\prod_{i=1}^{k_0} \theta_i^{m \hat{p}_i} (1-\theta_i)^{n \hat{q}_i} \phi_i^{m \hat{p}_i + n \hat{q}_i} \right] \cdot \left[a_{k_0}^{m'} b_{k_0}^{n'} \prod_{i=k_0+1}^{k} a_i^{m \hat{p}_i} b_i^{n \hat{q}_i} \right]. \tag{2.6}$$

Since the restrictions together with basic restrictions do not relate the first and second parts (each in bracket), the maximum of (2.6) can be obtained by maximizing two parts separately. The maximization of the second part is easy and hence we have $\hat{a}_{k_0} = m'/m$, $\hat{b}_{k_0} = n'/n$, and $a_i = \hat{p}_i$, $b_i = \hat{q}_i$ for $i = k_0+1, \cdots, k$.

The maximization of the first part is little bit complicated. The whole procedure is well described in Dykstra et al. (1995). Let $\bar{\phi}_i = m\hat{p}_i + n\hat{q}_i$ for $i = 1, \cdots, k_0$ and let $\bar{\theta} = E_w(\hat{\theta}|I)$ where $w = (m\hat{p}_1 + n\hat{q}_1, \cdots, m\hat{p}_{k_0} + n\hat{q}_{k_0})$. Then it follows from Theorem 1.3.3 of Robertson et al. (1988) that $\sum_{i=1}^{k_0} \bar{\theta}_i \bar{\phi}_i = m'$ and hence maximize (2.6) under (2.5) with basic restrictions. Evaluation of p and q at $\theta_i = \bar{\theta}_i$, $\phi_i = \bar{\phi}_i$ for $j = 1, \cdots, k-1$, and $a_i = \hat{a}_i$, $b_i = \hat{b}_i$ for $i = k_0, \cdots, k$ gives the restricted ML estimates $\bar{p}$ and $\bar{q}$.

Suppose $p = q$. Then $\tilde{p}_i = \tilde{q}_i = (m\hat{p}_i + n\hat{q}_i)/(m+m)$ for $i = 1, \cdots, k$. We have $\tilde{a}_{k_0} = \tilde{b}_{k_0} = (m'+n')/(m+n)$, $\tilde{a}_i = \tilde{b}_i = (m\hat{p}_i + n\hat{q}_i)/(m+m)$ for $i = k_0+1, \cdots, k$, and $\tilde{\theta}_i = m'/(m'+n')$, $\tilde{\phi}_i = (m\hat{p}_i + n\hat{q}_i)$ for $i = 1, \cdots, k_0$.

2.4 Local dominated ordering

Suppose $p_i \ge q_i$ for $i = 1, \cdots, k_0$. Let $a_i = p_i$, $b_i = q_i$ for $i = 1, \cdots, k_0$ and $a_{k_0+1} = \sum_{j=k_0+1}^{k} p_j$, $b_{k_0+1} = \sum_{j=k_0+1}^{k} q_j$. Let $\theta_i = p_i / a_{k_0+1}$ and $\eta_i = q_i / b_{k_0+1}$ for $i = k_0+1, \cdots, k$. Then the basic restriction become (1) $0 \le a_i, b_i \le 1$ for $i = 1, \cdots, k_0+1$, (2) $\sum_{i=1}^{k_0+1} a_i = \sum_{i=1}^{k_0+1} b_i = 1$, (3) $0 \le \theta_i, \eta_i \le 1$ for

$i=k_0+1,\cdots,k$, and (4) $\sum_{i=k_0+1}^{k}\theta_i=\sum_{i=k_0+1}^{k}\eta_i=1$. The local restriction becomes

$$a_i \geq b_i \text{ for } i=1,\cdots,k_0 \tag{2.7}$$

The likelihood function is proportional to

$$\left[\prod_{i=1}^{k_0} a_i^{m\hat{p}_i} b_i^{n\hat{q}_i} a_{k_0+1}^{m\sum_{\ell=k_0+1}^{k}\hat{p}_\ell} b_{k_0+1}^{n\sum_{\ell=k_0+1}^{k}\hat{q}_\ell}\right]\cdot\left[\prod_{i=k_0+1}^{k}\theta_i^{m\hat{p}_i}\eta_i^{n\hat{q}_i}\right]. \tag{2.8}$$

No restrictions relate the two parts (each in bracket) in (2.8) to each other and hence (2.8) can be maximized by maximizing two parts separately. From the second part we have $\bar{\theta}_i=\hat{\theta}_i=m\hat{p}_i/m'$, $\bar{\eta}_i=\hat{\eta}_i=n\hat{q}_i/n'$ for $i=k_0+1,\cdots,k$, where $m'=m\sum_{i=k_0+1}^{k}\hat{p}_i$ and $n'=n\sum_{i=k_0+1}^{k}\hat{q}_i$.

For the first part, Fenchel dual plays a fundamental role in estimation of restricted ML estimate of $\boldsymbol{a}=(a_1,\cdots,a_{k_0+1})$ and $\boldsymbol{b}=(b_1,\cdots,b_{k_0+1})$. The rigorous estimation procedure for the case of $k_0=k$ is given in Chang (1993). However we briefly state the estimation procedure here. Let $T=\{\boldsymbol{x}\in R^{k_0+1}: x_i\geq x_k, i=1,\cdots,k_0\}$. Then the Fenchel dual of T, denoted by T^{w*}, is given by

$$\{\boldsymbol{x}\in R^{k_0+1}: w_i x_i\leq 0,\ i=1,\cdots,k_0,\ \sum_{i=1}^{k_0+1} w_i x_i=0\}.$$

We also note that it follows from mathematical induction that

$$\bar{a}_i\geq\frac{m\hat{a}_i+n\hat{b}_i}{N}\geq\bar{b}_i \text{ for } i=1,\cdots,k_0. \tag{2.9}$$

Let B be defined by

$$\{\boldsymbol{u}\in R^{2k_0+2}: u_i\geq u_{k_0+1} \text{ for } i=1,\cdots,k_0, \text{ and } u_i\leq u_{2k_0+2} \text{ for } i=k_0+2,\cdots,2k_0+1\}.$$

Let

$$\begin{aligned} h_i &= N^{-1}+\frac{n}{mN}\frac{\hat{a}_i}{\hat{b}_i},\quad i=1,\cdots,k_0+1,\\ &= N^{-1}+\frac{m}{nN}\frac{\hat{a}_{i-k_0-1}}{\hat{b}_{i-k_0-1}},\quad i=k_0+2,\cdots,2k_0+2,\\ w &= (m\hat{a}_1,\cdots,m\hat{a}_{k_0+1},n\hat{b}_1,\cdots,n\hat{b}_{k_0+1}),\\ t &= (a_1/m\hat{a}_1,\cdots,a_{k_0+1}/m\hat{a}_{k_0+1},b_1/n\hat{b}_1,\cdots,b_{k_0+1}/n\hat{b}_{k_0+1}). \end{aligned}$$

Then restriction (2.9) together with basic restriction become

$$
\begin{aligned}
&(t_i-h_i)w_i \geq 0 \ \text{ for } \ i=1,\cdots,k_0, \\
&(t_i-h_i)w_i \leq 0 \ \text{ for } \ i=k_0+2,\cdots,2k_0+1, \ \text{ and} \\
&\sum_{i=1}^{k_0+1}(t_i-h_i)w_i = \sum_{i=k_0+2}^{2k_0+2}(t_i-h_i)w_i = 0.
\end{aligned}
\tag{2.10}
$$

Restriction (2.10) can be rewritten as $\mathbf{h}-\mathbf{t} \in B^{w*}$. Hence
$\bar{\mathbf{t}} = (\bar a_1/m\hat a_1,\cdots,\bar a_{k_0+1}/m\hat a_{k_0+1},\ \bar b_1/n\hat b_1,\cdots,\bar b_{k_0+1}/n\hat b_{k_0+1})$
solves

$$\text{minimize } \sum_{i=1}^{2k_0+2} w_i f(t_i) \ \text{ subject to } \ \mathbf{h}-\mathbf{t}\in B^{w*}.$$

Appealing to Theorem 3.4 of Barlow and Brunk (1972), we have $(\bar{a},\bar{b}) = \mathbf{w}E_{\mathbf{w}}(\mathbf{h}|B)$. Since membership in B imposes no restriction between the first k_0+1 coordinates and the last k_0+1 coordinates of a point, $(E(\cdot|B)_1,\cdots,E(\cdot|B)_{k_0+1})$ and $E(\cdot|B)_{k_0+2},\cdots,E(\cdot|B)_{2k_0+2})$ can be computed independently. It follows that

$$
\begin{aligned}
\bar{\mathbf{a}} &= \hat{\mathbf{a}}E_{\hat{\mathbf{a}}}\left(\frac{m\hat{\mathbf{a}}+n\hat{\mathbf{b}}}{N\hat{\mathbf{a}}}\Big|T\right),\\
\bar{\mathbf{b}} &= \hat{\mathbf{b}}E_{\hat{\mathbf{b}}}\left(\frac{m\hat{\mathbf{a}}+n\hat{\mathbf{b}}}{N\hat{\mathbf{b}}}\Big|T'\right) = -\hat{\mathbf{b}}E_{\hat{\mathbf{b}}}\left(-\frac{m\hat{\mathbf{a}}+n\hat{\mathbf{b}}}{N\hat{\mathbf{b}}}\Big|T\right),
\end{aligned}
$$

where $T' = \{\mathbf{x}\in R^{k_0+1} : -\mathbf{x}\in T\}$. Evaluation of $\mathbf{p}$ and $\mathbf{q}$ at $\mathbf{a}=\bar{\mathbf{a}}$, $\mathbf{b}=\bar{\mathbf{b}}$, $\theta_i=\bar\theta_i$, $\eta_i=\bar\eta_i$ for $i=1,\cdots,k_0$ gives the restricted ML estimates $\bar{\mathbf{p}}$ and $\bar{\mathbf{q}}$.

Suppose $\mathbf{p}=\mathbf{q}$. This is equivalent to $a_i=b_i$ for $i=1,\cdots,k_0+1$ and $\theta_i=\eta_i$ for $i=k_0+1,\cdots,k$.

$$
\begin{aligned}
\tilde a_i = \tilde b_i &= \frac{m\hat p_i + n\hat q_i}{N}, \quad \text{for } \ i=1,\cdots,k_0,\\
\tilde a_{k_0+1} = \tilde b_{k_0+1} &= \frac{m\sum_{j=k_0+1}^{k}\hat p_i + n\sum_{j=k_0+1}^{k}\hat q_i}{N},\\
\tilde\theta_i = \tilde\eta_i &= \frac{m\hat p_i + n\hat q_i}{\sum_{j=k_0+1}^{k}(m\hat p_j + n\hat q_j)}, \quad \text{for } \ i=k_0+1,\cdots,k.
\end{aligned}
$$

Finally we briefly discuss the strong consistency of the estimators given above. We note that all functions used for transformation of the parameter space are continuous with respect to their arguments. We also note that the projection operators $E_{\mathbf{w}}(\mathbf{x}|\cdot)$ used in this section are continuous with respect to $\mathbf{w}$ and $\mathbf{x}$. It follows from the strong law of large numbers together with these continuity properties of the functions and the projection operators that all restricted estimators preserve the strong consistency.

3. Likelihood Ratio Tests

In this section we consider the test of equality of two multinomial parameters against a local restriction. Let H_0: $p = q$ and let H_1 be corresponding local dependence restriction. The test for H_1 against all alternatives is not going to be discussed in this paper because it is basically the same as the test for the case that $k_0 = k$, i.e., full restriction. The asymptotic null distributions of each of four test statistics are chi-bar-square distributions. Each likelihood function is consisted of two parts; one is related to local restriction and the other is not. Moreover two parts do not relate to each other. This means that independence is guaranteed when we find the distribution of test statistics. The likelihood ratio statistic is also factored into two parts; one is related to local restriction and the other is not. After taking logarithm and multiplied by -2 on likelihood ratio, the first part has asymptotically a chi-bar-square distribution with corresponding partial order and weights. On the other hand, the second part has asymptotically chi-square distribution with certain degrees of freedom. Since the two parts are statistically independent, the asymptotic null distribution of test statistic is a convolution of two distributions, which is also a chi-bar-square distribution. Now we will see these for each of four cases.

3.1 Local stochastic ordering

The test rejects H_0 in favor of local stochastic ordering for the large value of

$$\begin{aligned} T_{01} = 2\Big[&\sum_{i=1}^{k_0} m\hat{p}_i(\ln \bar{a}_i - \tilde{a}_i) + \Big(\sum_{j=k_0+1}^{k} m\hat{p}_j\Big)(\ln \bar{a}_{k_0+1} - \ln \tilde{a}_{k_0+1}) \\ &+ \sum_{i=1}^{k_0} n\hat{q}_i(\ln \bar{b}_i - \tilde{b}_i) + \Big(\sum_{j=k_0+1}^{k} n\hat{q}_j\Big)(\ln \bar{b}_{k_0+1} - \ln \tilde{b}_{k_0+1})\Big] \\ &+ 2\sum_{i=k_0+1}^{k}[m\hat{p}_j(\ln \bar{\phi}_j - \ln \tilde{\phi}_j) + n\hat{q}_j(\ln \bar{\tau}_j - \ln \tilde{\tau}_j)]. \end{aligned}$$

Let T_{01}^1 be the first term and T_{01}^2 be the last term. It follows from Theorem 4.1 of Robertson and Wright (1981) that T_{01}^1 has a chi-bar square distribution, specifically, for every $t > 0$,

$$\lim_{m, n \to \infty} \Pr[T_{01}^1 \geq t] = \sum_{\ell=1}^{k_0+1} P_S(\ell, k_0+1; \tilde{a}) \Pr[\chi^2_{k_0+1-\ell} \geq t],$$

where $P_S(\ell, k_0+1; \tilde{a})$ is level probability with respect to simple ordering and weights $\tilde{a} = (\tilde{a}_1, \cdots, \tilde{a}_{k_0+1})$. And T^2_{01} has a chi-square distribution with $k-k_0-1$ degrees of freedom. Since T^1_{01} and T^2_{01} are independent, we have

$$\begin{aligned} \lim_{m,n\to\infty} \Pr[T_{01} \geq t] &= \sum_{\ell=1}^{k_0+1} P_S(\ell, k_0; \tilde{a}) \Pr[\chi^2_{k-\ell} \geq t] \\ &\leq \frac{1}{2}\Pr[\chi^2_{k-1} \geq t] + \frac{1}{2}\Pr[\chi^2_{k-2} \geq t]. \end{aligned} \tag{3.1}$$

Since $\tilde{a}$ is unknown, we need to use least favorable distribution, which is relate to (3.1), or plug in an estimate of $\tilde{a}$ to find a critical value. We will discuss about this later.

3.2 Local uniform stochastic ordering

The test rejects H_0 in favor of local uniform stochastic ordering for large value of T_{01} =

$$2\sum_{j=1}^{k-1}\Big[\big(\sum_{\ell=j+1}^{k} m\hat{p}_\ell\big)(\ln \bar{\theta}_{1j} - \ln \tilde{\theta}_{1j}) + m\hat{p}_j(\ln(1-\bar{\theta}_{1j}) - (1-\ln \tilde{\theta}_{1j})) + \big(\sum_{\ell=j+1}^{k} n\hat{q}_\ell\big)(\ln \bar{\theta}_{2j} - \ln \tilde{\theta}_{2j}) + n\hat{q}_j(\ln(1-\bar{\theta}_{2j}) - (1-\ln \tilde{\theta}_{2j}))\Big].$$

Let $T_{01} = \sum_{j=1}^{k-1} T^{(j)}_{01}$, say. We know that $T^{(j)}_{01}$ are statistically independent. Careful look at T_{01} reveals that each of $T^{(j)}_{01}$ for $j = 1, \cdots, k_0-1$ is likelihood ratio test statistic for testing equality of two binomial parameters against one-sided alternative, i.e., order restriction. Hence we have, for each $j = 1, \cdots, k_0-1$,

$$\lim_{m,n\to\infty} \Pr[T^{(j)}_{01} \geq t] = \frac{1}{2}\Pr[\chi^2_0 \geq t] + \frac{1}{2}\Pr[\chi^2_1 \geq t],$$

and since $T^{(j)}_{01}$ are independent we have

$$\lim_{m,n\to\infty} \Pr[\sum_{j=1}^{k_0-1} T^{(j)}_{01} \geq t] = \sum_{\ell=1}^{k_0} \binom{k_0-1}{\ell-1}\left(\frac{1}{2}\right)^{k_0-1} \Pr[\chi^2_{\ell-1} \geq t].$$

We also note that each of $T^{(j)}_{01}$ for $j = k_0, \cdots, k-1$ is likelihood ratio test statistic for testing equality of two binomial parameters against two-sided alternative and hence $\lim_{m,n\to\infty} \Pr[T^{(j)}_{01} \geq t] = \Pr[\chi^2_1 \geq t]$. By independence of $T^{(j)}_{01}$, we have

$$\lim_{m,n\to\infty} \Pr[T_{01} \geq t] = \sum_{\ell=1}^{k_0} \binom{k_0-1}{\ell-1}\left(\frac{1}{2}\right)^{k_0-1} \Pr[\chi^2_{k-k_0+\ell-1} \geq t].$$

It is of interest to observe that the asymptotic null distribution does not depend upon the unknown common value of two parameters.

3.3 Local likelihood ratio ordering

The test rejects H_0 in favor of local likelihood ratio ordering for large value of

$$\begin{aligned} T_{01} &= 2\sum_{i=1}^{k_0}\left[m\hat{p}_i(\ln\tilde{\theta}_i - \ln\bar{\theta}_i) + n\hat{q}_i(\ln(1-\tilde{\theta}_i) - (1-\ln\bar{\theta}_i))\right] \\ &\quad + 2\Big[m'(\ln\tilde{a}_{k_0} - \ln\bar{a}_{k_0}) + \sum_{i=k_0+1}^{k} m\hat{p}_i(\ln\tilde{a}_i - \ln\bar{a}_i) \\ &\quad + n'(\ln\tilde{b}_{k_0} - \ln\bar{b}_{k_0}) + \sum_{i=k_0+1}^{k} n\hat{q}_i(\ln\tilde{b}_i - \ln\bar{b}_i)\Big]. \end{aligned}$$

Let T_{01}^1 be the first term and T_{01}^2 be the last term. It follows from Theorem 3.1 of Dykstra {it et al.} (1995) that T_{01}^1 has a chi-bar square distribution, specifically, for every $t>0$,

$$\lim_{m,n\to\infty}\Pr[T_{01}^1\geq t] = \sum_{\ell=1}^{k_0} P_S(\ell, k_0;\ \tilde{w})\Pr[\chi^2_{\ell-1}\geq t],$$

where $P_S(\ell, k_0;\ \tilde{w})$ is level probability with respect to simple ordering and weights $\tilde{w}=(p_1,\cdots,p_{k_0})$. And T_{01}^2 has a chi-square distribution with $k-k_0$ degrees of freedom. Since T_{01}^1 and T_{01}^2 are independent, we have

$$\begin{aligned} \lim_{m,n\to\infty}\Pr[T_{01}\geq t] &= \sum_{\ell=1}^{k_0} P_S(\ell, k_0;\ \tilde{w})\Pr[\chi^2_{k-k_0+\ell-1}\geq t] \\ &\leq \sum_{\ell=1}^{k_0}\binom{k_0-1}{\ell-1}2^{-k_0+1}\Pr[\chi^2_{k-k_0+\ell-1}\geq t]. \end{aligned} \tag{3.2}$$

Since $\tilde{w}$ is unknown, we need to use least favorable distribution, which is relate to (3.2}), or plug in an estimate of $\tilde{w}$ to find a critical value. We will discuss about this later.

3.4 Local dominated ordering

The test rejects H_0 in favor of local dominated ordering for large value of

$$
\begin{aligned}
T_{01} = & \; 2\Big[\sum_{i=1}^{k_0} \{ m \hat{p}_i (\ln \tilde{a}_i - \ln \bar{a}_i) + n \hat{q}_i (\ln \tilde{b}_i - \ln \bar{b}_i) \} \\
& + (m \sum_{\ell = k_0+1}^{k} \hat{p}_\ell)(\ln \tilde{a}_{k_0+1} - \ln \bar{a}_{k_0+1}) \\
& + (n \sum_{\ell = k_0+1}^{k} \hat{q}_\ell)(\ln \tilde{b}_{k_0+1} - \ln \bar{b}_{k_0+1}) \Big] \\
& + 2\Big[\sum_{i=1}^{k_0} m \hat{p}_i (\ln \tilde{\theta}_i - \ln \bar{\theta}_i) + \sum_{i=1}^{k_0} n \hat{q}_i (\ln \tilde{\eta}_i - \ln \bar{\eta}_i) \Big].
\end{aligned}
$$

Let T_{01}^1 be the first term and T_{01}^2 be the last term. It follows from Theorem 2.2.5 of Chang (1993) that T_{01}^1 has a chi-bar square distribution, specifically, for every $t > 0$,

$$
\lim_{m,n\to\infty} \Pr[T_{01}^1 \ge t] = \sum_{\ell=1}^{k_0+1} P_T(\ell, k_0+1; \tilde{a}) \Pr[\chi^2_{k_0+1-\ell} \ge t],
$$

where $P_T(\ell, k_0+1; \tilde{a})$ is level probability with respect to simple tree ordering and weights $\tilde{a} = (\tilde{a}_1, \cdots, \tilde{a}_{k_0+1})$. And T_{01}^2 has a chi-square distribution with $k - k_0 - 1$ degrees of freedom. Since T_{01}^1 and T_{01}^2 are independent, we have

$$
\begin{aligned}
\lim_{m,n\to\infty} \Pr[T_{01} \ge t] &= \sum_{\ell=1}^{k_0+1} P_S(\ell, k_0; \tilde{a}) \Pr[\chi^2_{k-\ell} \ge t] \\
&\le \frac{1}{2} \Pr[\chi^2_{k-1} \ge t] + \frac{1}{2} \Pr[\chi^2_{k-2} \ge t].
\end{aligned} \tag{3.3}
$$

Since $\tilde{a}$ is unknown, we need to use least favorable distribution, which is related to (3.3), or plug in an estimate of $\tilde{a}$ to find a critical value. We will discuss about this next.

The asymptotic null distributions except for local uniform stochastic ordering case depend upon unknown common parameter through transformed parameters. Note $p = q$. The asymptotic null distributions for local stochastic, local likelihood ratio and local dominated orderings depend on $\tilde{a} = (p_1, \cdots, p_{k_0}, \sum_{j=k_0+1}^{k} p_j)$, $\tilde{w} = (p_1, \cdots, p_{k_0})$, and $\tilde{a} = (p_1, \cdots, p_{k_0}, \sum_{j=k_0+1}^{k} p_j)$, respectively. Since these quantities are unknown the critical values are intractable. The frequently used method is to use least favorable distribution, which is the stochastically largest distribution over the parameter space satisfying null hypothesis. This test, however, very conservative for some values of parameter satisfying alternative hypothesis. Another frequently used method is approximation of asymptotic null distribution. This approximation is obtained by computing level probability approximately. There are several algorithm for approximating level probabilities. One is equal-weights level probability. The level probability, which depend heavily upon unknown quantities, are known to be robust with respect to weights. This fact lead us to use equal-weight level probability which is easy to compute for some well known orderings such as simple and simple tree orderings. For local likelihood ratio ordering case this will work reasonably well. But for local stochastic ordering and dominated ordering this won't work well if $\sum_{j=k_0+1}^{k} p_j$ is relatively large than p_i, $i=1,\cdots,k_0$, which is likely to happen. Another method is to plug in an estimate of weights and compute level probability. This is valid only if the dimension of weight vector is equal to or less than 5. Robertson and Wright (1983) studied approximation of null distribution by so-called pattern approximation. A FORTRAN program for computing level probability with respect to simple order and arbitrary weights is given in Pillers, Robertson, and Wright (1984). See also Cran (1981).

4. Examples

To illustrate the inferential procedures discussed in earlier sections, here we analyze the data set discussed in introduction. Table ref{table1} shows the numbers of admitted students to the college of oriental studies in 1997 and 1998 for each of 10 groups which classify the students according to their high school ranks. The students with high school rank 5 or higher are considered to have good academic qualification. Apparently, the proportion of well qualified students in 1998 is larger than that in 1997. This, however, does not mean that the admitted students in 1998 are likely to be well qualified than that in 1997. If it is true then we must see that the proportion of students with rank 1 in 1998 is larger than in 1997, and the proportion of students with rank 1 and 2 is larger than in 1997, and so on. This trend is associated with stochastic ordering. Since we focussed on the students with rank 5 or higher, it becomes local dependency rather than usual stochastic ordering. Specifically, using the same notation given in introduction, we are going to test $H_0: p_i = q_i$ for $i=1,\cdots,10$ against $H_1: \sum_{j=1}^{i} p_j \geq \sum_{j=1}^{i} q_j$ for $i=1,\cdots,5$ Note that we set $k_0=5$.

Table 4.1 Students admitted to the College of Oriental Studies, Pusan University of Foreign Studies in the academic year 1997 and 1998. Classified into 10 groups according to their high school rank.

year	High School Rank										
	1	2	3	4	5	6	7	8	9	10	total
1997	4	10	12	11	44	57	81	37	15	9	280
1998	4	6	18	31	41	53	38	15	0	2	208

The computational details for estimating cell frequencies under H_0 and H_1 are shown in Table 4.2. The computed value of the likelihood ratio statistic T_{01} is 52.8349. We computed estimated level probability using the A FORTRAN program given in Pillers et al. (1984) and Cran(1981). These probabilities are 0.2071, 0.4108, 0.2793, 0.0884, 0.0135, and 0.0008 for $\ell = 1, \cdots, 6$, respectively. The p-value is 1.25086×10^{-8} and the test rejects the null hypothesis. The level probabilities with respect to equal-weights are 0.16667, 0.38056, 0.31250, 0.11806, 0.02083, and 0.00139 for $\ell = 1, \cdots, 6$, respectively. The p-value based on equal-weights level probabilities is 1.10593×10^{-8}. We omit the analysis result for other types of local restriction.

Table 4.2 Computational details for estimation of cell probabilities under local stochastic ordering.

	a_i	b_i	ϕ_i	τ_i	$\overline{a}_i$	$\overline{b}_i$	$\overline{p}_i$	$\overline{q}_i$	$\overline{p}_i = \overline{q}_i$
1	0.019	0.014	-	-	0.020	0.014	0.020	0.014	0.016
2	0.029	0.036	-	-	0.030	0.035	0.030	0.035	0.030
3	0.087	0.043	-	-	0.086	0.043	0.086	0.043	0.061
4	0.149	0.039	-	-	0.149	0.039	0.149	0.039	0.086
5	0.197	0.157	-	-	0.197	0.157	0.197	0.157	0.174
6	0.519	0.711	0.491	0.286	0.519	0.711	0.255	0.204	0.225
7	-	-	0.352	0.407	-	-	0.182	0.290	0.244
8	-	-	0.139	0.186	-	-	0.072	0.132	0.107
9	-	-	0.000	0.075	-	-	0.000	0.054	0.031
10	-	-	0.019	0.045	-	-	0.010	0.032	0.023

Recall that the test statistic is consisted of two parts. One is related to local restriction and the other is not. For our example the computed value of test statistic is sum of 28.9580 and 23.8769. Suppose that test reject the null hypothesis but the first part is relatively small compare to the second part. For this case one can not claim that the local restriction contribute significantly to the test statistic. From this we can conclude that the tests suggested in this paper are not recommendable for such cases.

Note that two parts in test statistics are statistically independent. Let F be the ratio of first part to second part. For local stochastic ordering, $F = T^1_{01}/T^2_{01}$. Under H_0, the distribution of F is mixture of F distributions. Specifically, for local stochastic ordering,

$$\lim_{m,n\to\infty} \Pr[F \geq t] = \sum_{\ell=1}^{k_0+1} P_S(\ell, k_0+1; \tilde{a}) \Pr[F_{(k_0+1-\ell, k-k_0-1)} \geq t],$$

where $F_{(\nu_1, \nu_2)}$ is F-distributed variables with degrees of freedom ν_1 and ν_2. From this we can test how the local restriction contribute to the test statistic. We note this test can not be used for testing local restriction.

참고문헌

1. Barlow, R. E. and Brunk, H. D. (1972). The isotonic regression problem and its dual. *Journal of the American Statistical Association*, 67, 140-147.
2. Chang, S. (1993). *Statistical Inferences of Multinomial Distributions with Order Restrictions*, Ph. D. Thesis, The University of Iowa.
3. Cran, G. W. (1981). Calculation of the probability $P(\ell, k)$ for the simply ordered alternative, *Journal of the Royal Statistical Society (C)*, 30, 85-91.
4. Douglas, R., S. E. Fienberg, M. T. Lee, A. R. Sampson, and L. R. Whitaker (1990). Positive dependence concepts for ordinal contingency tables, In *Topics in Statistical Dependence* (H. W. Block, A. R. Sampson, T. H. Savits eds.). IMS Lecture/Monograph series, Hayward, CA, 189-202.
5. Dykstra, R. L., S. C. Kochar, and T. Robertson (1991). Statistical inference for uniform stochastic ordering in several populations, *The Annals of Statistics*, 19, 870-888.
6. 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.
7. Park, C. G., Lee, C. I. C., and Robertson, T. (1998). Goodness-of-fit test for uniform stochastic ordering among several distributions. *The*

Canadian Journal of Statistics, 26, 69-81.

8. Pillers, C., Robertson T., and Wright F. T. (1984). A FORTRAN program for the level probabilities of ordered restricted inference, *Journal of the Royal Statistical Society (C)*, 33, 115-119.
9. Robertson, T. and Wright, F. T. (1981). Likelihood ratio tests for and against a stochastic ordering between multinomial populations, *The Annals of Statistics,* 9, 1248-1257.
10. Robertson, T., Wright, F. T., and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*, Wiley, Chichester.

[received date : Feb. 2003, accepted date : May. 2003]