

On Asymptotic Properties of a Maximum Likelihood Estimator of Stochastically Ordered Distribution Function

Myongsik Oh^{1,a}

^aDepartment of Statistics, Busan University of Foreign Studies

Abstract

Kiefer (1961) studied asymptotic behavior of empirical distribution using the law of the iterated logarithm. Robertson and Wright (1974a) discussed whether this type of result would hold for a maximum likelihood estimator of a stochastically ordered distribution function; however, we show that this cannot be achieved. We provide only a partial answer to this problem. The result is applicable to both estimation and testing problems under the restriction of stochastic ordering.

Keywords: Law of the iterated logarithm, maximum likelihood estimation, stochastically ordered distribution function.

1. Introduction

Kiefer (1961) showed that if a distribution function F is absolutely continuous with respect to a Lebesgue measure, then

$$P \left[\lim_{n \rightarrow \infty} \left(\frac{n}{\ln \ln n} \right)^{\frac{1}{2}} \sup_x |\hat{F}_n(x) - F(x)| = \sqrt{2} \right] = 1,$$

where \hat{F}_n is an empirical distribution of F . Robertson and Wright (1974a) discussed about if such an iterated logarithmic results holds for a maximum likelihood estimator of distribution functions under stochastic ordering.

This paper investigates if the constrained estimator of multinomial parameter, p_i , under stochastic ordering has a Kiefer type asymptotic property, *i.e.*, the asymptotic behavior of $\max_{1 \leq i \leq k} |\bar{p}_i^* - p_i|$, where \bar{p}_i^* is a constrained estimator of a multinomial parameter under stochastic ordering.

In a multinomial setting with parameter $\mathbf{p} = (p_1, p_2, \dots, p_k)$, a maximum likelihood estimator, \hat{p}_i^* , of p_i under a certain order restriction satisfies the following properties;

$$P \left[\limsup_{n \rightarrow \infty} \left(\frac{n}{\ln \ln n} \right)^{\frac{1}{2}} \max_{1 \leq i \leq k} |\hat{p}_i^* - p_i| = \max_{1 \leq i \leq k} [2p_i(1 - p_i)]^{\frac{1}{2}} \right] = 1.$$

The proof of this result can be seen in Robertson *et al.* (1988). Part of this proof depends on the property (reduction of error) *i.e.*,

$$\max_{1 \leq i \leq k} |\hat{p}_i^* - p_i| \leq \max_{1 \leq i \leq k} |\hat{p}_i - p_i|.$$

This work was supported by the research grant of the Busan University of Foreign Studies in 2011.

¹ Professor, Department of Statistics, Busan University of Foreign Studies, 15 SukPo-Ro, Nam-Gu, Busan 608-738, Korea.
E-mail: moh@pufs.ac.kr

See Robertson and Wright (1974b) for details.

However this type of property does not hold for stochastic ordering problem as can be seen in the following. Suppose p is observed to be (0.35, 0.11, 0.18, 0.12, 0.24) and q is (0.2, 0.2, 0.2, 0.2, 0.2). Then the constrained MLE of p is

$$\frac{1}{380}(140, 44, 72, 48, 76),$$

and hence

$$0.16842 = \max_{1 \leq i \leq 5} |\hat{p}^* - q_i| \geq \max_{1 \leq i \leq 5} |\hat{p} - q_i| = 0.15.$$

This means that we cannot expect that the conjecture given by Robertson and Wright (1974) will come true. However we are able to give a partial answer to this problem. This fact is quite useful in some testing problems under stochastic ordering.

2. Some Issues in Estimation and Test

In this section, we discuss the one-sample problem for a stochastic ordering between two multinomial parameters.

Now let $\mathbf{p} = (p_1, p_2, \dots, p_k)$ and $\mathbf{q} = (q_1, q_2, \dots, q_k)$ be two multinomial parameters. We assume that both \mathbf{p} and \mathbf{q} are in $\{\mathbf{x} \in \mathbf{R}^k : x_i > 0, \sum_{i=1}^k x_i = 1\}$ and \mathbf{q} is known. The stochastic ordering between \mathbf{p} and \mathbf{q} can be expressed as

$$\sum_{j=1}^i p_j \geq \sum_{j=1}^i q_j, \quad i = 1, 2, \dots, k-1, \quad \sum_{j=1}^k p_j = \sum_{j=1}^k q_j.$$

Now let $\hat{\mathbf{p}}$ be the vector of relative frequencies of a sample of size m from the \mathbf{p} population. Robertson and Wright (1981) provides the restricted MLE of \mathbf{p} under H_1 as follows; If $\hat{p}_i > 0, i = 1, 2, \dots, k$, then the restricted MLE, $\bar{\mathbf{p}}$, of \mathbf{p} is

$$\bar{\mathbf{p}} = \hat{\mathbf{p}} E_{\hat{\mathbf{p}}} \left(\frac{\mathbf{q}}{\hat{\mathbf{p}}} \mid C \right).$$

where $C = \{\mathbf{x} \in \mathbf{R}^k : x_1 \geq x_2 \geq \dots \geq x_k\}$ and, for $\mathbf{x}, \mathbf{y} \in \mathbf{R}^k$, \mathbf{xy} denotes the vector $(x_1 y_1, x_2 y_2, \dots, x_k y_k)$ and $\mathbf{x}/\mathbf{y} = (x_1/y_1, \dots, x_k/y_k)$. They also prove that $P\{\lim_{m \rightarrow \infty} \bar{\mathbf{p}} = \mathbf{p}\} = 1$ provided $\mathbf{p} \gg \mathbf{q}$. If some of \hat{p}_i 's are equal to zero, the restricted MLE cannot be obtained by the above. See Lee (1987) for this case. It provides a restricted MLE of \mathbf{p} when some of \hat{p}_i 's are zero.

Suppose $\mathbf{p} \gg \mathbf{q}$. Let $D_{\mathbf{pq}} = \{\eta_1, \eta_2, \dots, \eta_A\}$ with $0 = \eta_0 < \eta_1 < \dots < \eta_A = k$ and

$$\begin{aligned} p_1 + \dots + p_i &= q_1 + \dots + q_i, & \text{for } i = \eta_1, \eta_2, \dots, \eta_A, \\ p_1 + \dots + p_i &> q_1 + \dots + q_i, & \text{for } i \neq \eta_1, \eta_2, \dots, \eta_A. \end{aligned}$$

Note that $D_{\mathbf{pq}}$ is nonempty. This is quite important to explain the asymptotic behavior in the estimation and testing problem. We briefly state the application to the testing problem. Consider the likelihood ratio test procedure to test stochastic ordering against all alternatives. Let H_1 be the hypothesis associated to stochastic ordering and H_2 be all alternatives. Robertson and Wright (1981) studied the

likelihood ratio test and gave the limiting distribution of the test statistic under H_1 . Following their notation the test rejects H_1 for large value of

$$S_{12} = -2m \sum_{i=1}^k \hat{p}_i (\ln \bar{p}_i - \ln \hat{p}_i). \quad (2.1)$$

Then for all t , $\lim_{m \rightarrow \infty} P[S_{12} \geq t] = \sum_{\ell=1}^A P(\ell, A; \mathbf{q}^+) P[\chi_{\ell-1}^2 \geq t]$, where $\mathbf{q}^+ = (q_1^+, q_2^+, \dots, q_A^+)$, $q_i^+ = \sum_{j=\eta_{i-1}+1}^{\eta_i} q_j$, $i = 1, 2, \dots, A$. The limiting null distribution depends on \mathbf{p} through A , the η_i 's and \mathbf{q}^+ . To approximate the null distribution one needs to estimate A , η_i 's and \mathbf{q}^+ . Since A and \mathbf{q}^+ are determined according to η_i 's one only needs to estimate η_i 's.

3. The Main Result

Theorem 1. For each $\eta \in D_{\mathbf{pq}}$,

$$P \left[\limsup_{m \rightarrow \infty} \left(\frac{m}{\ln \ln m} \right)^{\frac{1}{2}} \left| \sum_{j=1}^{\eta} \bar{p}_j - \sum_{j=1}^{\eta} q_j \right| \leq \sqrt{2} \right] = 1.$$

To prove theorem we need the following two lemmas. Before we mention the lemmas we briefly describe the computation of $E_{\mathbf{w}}(\mathbf{x}|C)$. For S , a nonempty subset of $\{1, 2, \dots, k\}$, set

$$Av(S) = \sum_{i \in S} w_i x_i / \sum_{i \in S} w_i.$$

Set $i_0 = 0$ and choose i_1 the largest positive integer i that maximizes $Av(\{i_0 + 1, \dots, i\})$. Next choose i_2 the largest integer i greater than i_1 that maximizes $Av(\{i_1 + 1, \dots, i\})$. Continuing this process, we obtain $0 = i_0 < i_1 < \dots < i_\ell = k$ and the projection

$$E_{\mathbf{w}}(\mathbf{x}|C)_i = Av(\{i_{j-1} + 1, \dots, i_j\}), \quad \text{for } i \in \{i_{j-1} + 1, \dots, i_j\} \text{ and } j = 1, 2, \dots, \ell.$$

The sets $\{i_{j-1} + 1, \dots, i_j\}$ are called the level sets. Details regarding level sets are discussed in Robertson *et al.* (1988).

Lemma 1 provides the general form of the level sets to compute $E_{\hat{\mathbf{p}}}(\mathbf{q}/\hat{\mathbf{p}}|C)$ when the sample size m is sufficiently large.

Lemma 1. For almost all ω (in the underlying probability space) there exists an $m_0(\omega)$ such that if $m \geq m_0(\omega)$ then the level sets to compute the projection $E_{\hat{\mathbf{p}}}(\mathbf{q}/\hat{\mathbf{p}}|C)$ are of the form $\{\eta_j + 1, \dots, \eta_\ell\}$ with $0 \leq j < \ell \leq A$.

Lemma 2. Suppose $i_0 \in D_{\mathbf{pq}}$ so that there is an ℓ_0 such that $i_0 = \eta_{\ell_0}$ and $1 \leq \ell_0 \leq A$. Then for almost all ω (in the underlying probability space) there exists an $m_0(\omega)$ such that if $m \geq m_0(\omega)$ then

$$\left| \sum_{j=1}^{i_0} \bar{p}_j - \sum_{j=1}^{i_0} q_j \right| \leq \max_{0 \leq \alpha < \ell_0 \leq \beta \leq A} \left| \sum_{j=\eta_\alpha+1}^{\eta_\beta} \hat{p}_j - \sum_{j=\eta_\alpha+1}^{\eta_\beta} p_j \right| + \max_{0 \leq \alpha < \ell_0} \left| \sum_{j=\eta_\alpha+1}^{i_0} \hat{p}_j - \sum_{j=\eta_\alpha+1}^{i_0} p_j \right|.$$

Proof of Theorem 1: First we assume that $i_0 \in D_{\mathbf{pq}}$ so that there exists an ℓ_0 such that $i_0 = \eta_{\ell_0}$ and $1 \leq \ell_0 \leq A$. Then by Lemma 2, there exists a set E_1 such that $P(E_1) = 1$ and $\omega \in E_1$ implies that there

exists an $m_0(\omega)$ such that $m \geq m_0(\omega)$ implies that

$$\left| \sum_{j=1}^{i_0} \bar{p}_j - \sum_{j=1}^{i_0} q_j \right| \leq \max_{0 \leq \alpha < \ell_0 \leq \beta \leq A} \left| \sum_{j=\eta_\alpha+1}^{\eta_\beta} \hat{p}_j - \sum_{j=\eta_\alpha+1}^{\eta_\beta} p_j \right| + \max_{0 \leq \alpha < \ell_0} \left| \sum_{j=\eta_\alpha+1}^{i_0} \hat{p}_j - \sum_{j=\eta_\alpha+1}^{i_0} p_j \right|. \quad (3.1)$$

By multiplying both sides by $\sqrt{m/\ln \ln m}$ and taking $\limsup_{m \rightarrow \infty}$ on both sides of (3.1) we have

$$\begin{aligned} \limsup_{m \rightarrow \infty} \left(\frac{m}{\ln \ln m} \right)^{\frac{1}{2}} \left| \sum_{j=1}^{i_0} \bar{p}_j - \sum_{j=1}^{i_0} q_j \right| &\leq \max_{0 \leq \alpha < \ell_0 \leq \beta \leq A} \limsup_{m \rightarrow \infty} \left(\frac{m}{\ln \ln m} \right)^{\frac{1}{2}} \left| \sum_{j=\eta_\alpha+1}^{\eta_\beta} \hat{p}_j - \sum_{j=\eta_\alpha+1}^{\eta_\beta} p_j \right| \\ &\quad + \max_{0 \leq \alpha < \ell_0} \limsup_{m \rightarrow \infty} \left(\frac{m}{\ln \ln m} \right)^{\frac{1}{2}} \left| \sum_{j=\eta_\alpha+1}^{i_0} \hat{p}_j - \sum_{j=\eta_\alpha+1}^{i_0} p_j \right|. \end{aligned} \quad (3.2)$$

The inequality is because any real sequences $\{a_n\}$ and $\{b_n\}$ $\limsup(a_n + b_n) \leq \limsup a_n + \limsup b_n$ and $\limsup(a_n \vee b_n) \leq \limsup a_n \vee \limsup b_n$, where \vee denotes the larger of the two numbers. See Section 4 of Billingsley (1986) for details.

Now by Kolmogorov's law of the iterated logarithm, for each α and β such that $0 \leq \alpha < \ell_0 \leq \beta \leq A$ and i_0 , we have, with probability one,

$$\limsup_{m \rightarrow \infty} \sqrt{\frac{m}{\ln \ln m}} \left| \sum_{j=\eta_\alpha+1}^{\eta_\beta} \hat{p}_j - \sum_{j=\eta_\alpha+1}^{\eta_\beta} p_j \right| = \sqrt{2 \sum_{j=\eta_\alpha+1}^{\eta_\beta} p_j \left(1 - \sum_{j=\eta_\alpha+1}^{\eta_\beta} p_j \right)}, \quad (3.3)$$

$$\limsup_{m \rightarrow \infty} \sqrt{\frac{m}{\ln \ln m}} \left| \sum_{j=\eta_\alpha+1}^{i_0} \hat{p}_j - \sum_{j=\eta_\alpha+1}^{i_0} p_j \right| = \sqrt{2 \sum_{j=\eta_\alpha+1}^{i_0} p_j \left(1 - \sum_{j=\eta_\alpha+1}^{i_0} p_j \right)}. \quad (3.4)$$

We may assume that (3.3) and (3.4) are true for such ω . Hence (3.2) becomes

$$\begin{aligned} &\limsup_{m \rightarrow \infty} \left(\frac{m}{\ln \ln m} \right)^{\frac{1}{2}} \left| \sum_{j=1}^{i_0} \bar{p}_j - \sum_{j=1}^{i_0} q_j \right| \\ &\leq \max_{0 \leq \alpha < \ell_0 \leq \beta \leq A} \sqrt{2 \sum_{j=\eta_\alpha+1}^{\eta_\beta} p_j \left(1 - \sum_{j=\eta_\alpha+1}^{\eta_\beta} p_j \right)} + \max_{0 \leq \alpha < \ell_0} \sqrt{2 \sum_{j=\eta_\alpha+1}^{i_0} p_j \left(1 - \sum_{j=\eta_\alpha+1}^{i_0} p_j \right)} \\ &\leq 2 \sqrt{\frac{1}{2}} = \sqrt{2}. \end{aligned}$$

The last inequality follows from the fact that $2p(1-p) \leq 1/2$ for $0 \leq p \leq 1$. \square

4. Remarks

For the case of two-sample problem, *i.e.*, both \mathbf{p} and \mathbf{q} are unknown, we can also obtain the similar result as Theorem 1.

Using Theorem 1 of previous section we can find a strongly consistent estimator of $D_{\mathbf{pq}}$. An example of such estimator is

$$D_{\mathbf{pq}}(\mathbf{c}) = \left\{ i \in (1, 2, \dots, k) : \sum_{j=1}^i \bar{p}_j - \sum_{j=1}^i q_j \leq c_i \right\}$$

for suitable choice of $\mathbf{c} = (c_1, c_2, \dots, c_k), c_i > 0$. Note that $D_{\mathbf{pq}}(\mathbf{c})$ is nonempty because it contains k . Based on this fact we can approximate the asymptotic distribution of S_{12} . Details regarding this approximation will appear elsewhere.

Appendix: Proofs

Proof of Lemma 1: By the strong law of large numbers there exists a set E such that $P(E) = 1$ and $\omega \in E$ implies that there exists an $m_0(\omega)$ and $\epsilon > 0$ for which

$$\frac{q_{\eta_j+1} + \dots + q_i}{\hat{p}_{\eta_j+1} + \dots + \hat{p}_i} < 1 - \epsilon, \quad (\text{A.1})$$

for each $j = 0, \dots, A-1$ and $i > \eta_j$ with $i \neq \eta_{j+1}, \dots, \eta_A$, and

$$\frac{q_{\eta_j+1} + \dots + q_{\eta_\ell}}{\hat{p}_{\eta_j+1} + \dots + \hat{p}_{\eta_\ell}} > 1 - \epsilon, \quad (\text{A.2})$$

for each $0 \leq j < \ell \leq A$ provided $m \geq m_0(\omega)$.

Now consider the pool adjacent violators algorithm(PAVA) in computing the projection $E_{\hat{\mathbf{p}}}(\mathbf{q}|\hat{\mathbf{p}}|C)$ for such m and ω . By (A.1) and (A.2), for each $j = 0, \dots, A-1$ and $\eta_j + 1 \leq i < \eta_{j+1}$,

$$\frac{q_{\eta_j+1} + \dots + q_i}{\hat{p}_{\eta_j+1} + \dots + \hat{p}_i} < 1 - \epsilon < \frac{q_{\eta_j+1} + \dots + q_{\eta_{j+1}}}{\hat{p}_{\eta_j+1} + \dots + \hat{p}_{\eta_{j+1}}}.$$

By the strict Cauchy mean value function property, which is shown in Robertson and Wright (1974b), this implies

$$\frac{q_{\eta_j+1} + \dots + q_i}{\hat{p}_{\eta_j+1} + \dots + \hat{p}_i} < \frac{q_{i+1} + \dots + q_{\eta_{j+1}}}{\hat{p}_{i+1} + \dots + \hat{p}_{\eta_{j+1}}},$$

and

$$\frac{q_{\eta_j+1} + \dots + q_i}{\hat{p}_{\eta_j+1} + \dots + \hat{p}_i} < 1 - \epsilon < \frac{q_{\eta_j+1} + \dots + q_{\eta_{j+1}}}{\hat{p}_{\eta_j+1} + \dots + \hat{p}_{\eta_{j+1}}} < \frac{q_{i+1} + \dots + q_{\eta_{j+1}}}{\hat{p}_{i+1} + \dots + \hat{p}_{\eta_{j+1}}}. \quad (\text{A.3})$$

By (A.3) and (A.1) we have

$$\frac{q_{\eta_j+1} + \dots + q_{i_1}}{\hat{p}_{\eta_j+1} + \dots + \hat{p}_{i_1}} < 1 - \epsilon < \frac{q_{i_2} + \dots + q_{\eta_{j+1}}}{\hat{p}_{i_2} + \dots + \hat{p}_{\eta_{j+1}}},$$

for $\eta_j + 1 \leq i_1 < \eta_{j+1}$ and $\eta_j \leq i_2 \leq \eta_{j+1}$. Hence in using the PAVA to compute the projection, $q_{\eta_j+1}/\hat{p}_{\eta_j+1}, \dots, q_{\eta_{j+1}}/\hat{p}_{\eta_{j+1}}$ will be pooled.

Now the projection can be obtained as follows.

$$E_{\hat{\mathbf{p}}} \left(\frac{\mathbf{q}}{\hat{\mathbf{p}}} \middle| C \right)_i = E_{\hat{\mathbf{p}}^+} \left(\frac{\mathbf{q}^+}{\hat{\mathbf{p}}^+} \middle| C^+ \right)_j$$

for $j = 0, \dots, A-1$ and $\eta_j + 1 \leq i \leq \eta_{j+1}$, where

$$\begin{aligned} \hat{\mathbf{p}}^+ &= \left(\sum_{j=\eta_0+1}^{\eta_1} \hat{p}_j, \dots, \sum_{j=\eta_{A-1}+1}^{\eta_A} \hat{p}_j \right), \quad \text{and} \\ \frac{\mathbf{q}^+}{\hat{\mathbf{p}}^+} &= \left(\sum_{j=\eta_0+1}^{\eta_1} q_j / \sum_{j=\eta_0+1}^{\eta_1} \hat{p}_j, \dots, \sum_{j=\eta_{A-1}+1}^{\eta_A} q_j / \sum_{j=\eta_{A-1}+1}^{\eta_A} \hat{p}_j \right), \end{aligned}$$

and $C^+ = \{\mathbf{x} \in \mathbf{R}^A : x_1 \geq x_2 \geq \dots \geq x_A\}$. Therefore the level sets are of the form $\{\eta_j + 1, \dots, \eta_\ell\}$ with $0 \leq j < \ell \leq A$. This completes the proof. \square

Proof of Lemma 2: By Lemma 1, there exists a set E such that $P(E) = 1$ and $\omega \in E$ implies that there exists an $m_0(\omega)$ such that $m \geq m_0(\omega)$ implies that the level sets in computing $E_{\hat{\mathbf{p}}}(\mathbf{q}|\hat{\mathbf{p}}|C)$ are of the form $\{\eta_j + 1, \dots, \eta_\ell\}$ with $0 \leq j < \ell \leq A$.

Now we fix ω and $m \geq m_0(\omega)$. Suppose the level sets for such ω and m are $\{\xi_\ell + 1, \dots, \xi_{\ell+1}\}$ for $\ell = 0, \dots, L \leq A - 1$ with $\xi_0 = 0$ and $\xi_{L+1} = \eta_A = k$. Note that $\{\xi_1, \dots, \xi_{L+1}\} \subseteq \{\eta_1, \eta_2, \dots, \eta_k\}$ and ξ 's depend on ω as well as m . Then for $i = 1, 2, \dots, k$ and $\ell = 0, \dots, L$

$$E_{\hat{\mathbf{p}}} \left(\frac{\mathbf{q}}{\hat{\mathbf{p}}} \middle| C \right)_i = \frac{\sum_{j=\xi_\ell+1}^{\xi_{\ell+1}} q_j}{\sum_{j=\xi_\ell+1}^{\xi_{\ell+1}} \hat{p}_j} \quad \text{with } \xi_\ell + 1 \leq i \leq \xi_{\ell+1}.$$

Hence we have

$$\bar{p}_i = \hat{p}_i \sum_{j=\xi_\ell+1}^{\xi_{\ell+1}} q_j \middle/ \sum_{j=\xi_\ell+1}^{\xi_{\ell+1}} \hat{p}_j \quad \text{with } \xi_\ell + 1 \leq i \leq \xi_{\ell+1}.$$

Consider a level set containing i_0 . Since $i_0 \in D_{\mathbf{pq}} = \{\eta_1, \eta_2, \dots, \eta_k\}$, we can choose an ℓ_1 such that $\xi_{\ell_1} + 1 \leq i_0 \leq \xi_{\ell_1+1}$, i.e., the level set containing i_0 is $\{\xi_{\ell_1} + 1, \dots, \xi_{\ell_1+1}\}$. Then we have

$$\begin{aligned} \sum_{j=1}^{i_0} \bar{p}_j - \sum_{j=1}^{i_0} q_j &= \sum_{\ell=0}^{\ell_1-1} \left(\sum_{j=\xi_\ell+1}^{\xi_{\ell+1}} \bar{p}_j - \sum_{j=\xi_\ell+1}^{\xi_{\ell+1}} q_j \right) + \sum_{j=\xi_{\ell_1}+1}^{i_0} \bar{p}_j - \sum_{j=\xi_{\ell_1}+1}^{i_0} q_j \\ &= \sum_{\ell=0}^{\ell_1-1} \left(\sum_{j=\xi_\ell+1}^{\xi_{\ell+1}} \hat{p}_j \frac{\sum_{j=\xi_\ell+1}^{\xi_{\ell+1}} q_j}{\sum_{j=\xi_\ell+1}^{\xi_{\ell+1}} \hat{p}_j} - \sum_{j=\xi_\ell+1}^{\xi_{\ell+1}} q_j \right) + \sum_{j=\xi_{\ell_1}+1}^{i_0} \bar{p}_j - \sum_{j=\xi_{\ell_1}+1}^{i_0} q_j \\ &= \sum_{j=\xi_{\ell_1}+1}^{i_0} \bar{p}_j - \sum_{j=\xi_{\ell_1}+1}^{i_0} q_j. \end{aligned}$$

Now we are going to find the upper bound of fluctuations when the sequence $\{\sum_{j=1}^{i_0} \bar{p}_j\}$ converges to $\sum_{j=1}^{i_0} q_j$ as the sample size m approaches infinity.

$$\begin{aligned} \left| \sum_{j=1}^{i_0} \bar{p}_j - \sum_{j=1}^{i_0} q_j \right| &= \left| \sum_{j=\xi_{\ell_1}+1}^{i_0} \bar{p}_j - \sum_{j=\xi_{\ell_1}+1}^{i_0} q_j \right| = \left| \sum_{j=\xi_{\ell_1}+1}^{i_0} \bar{p}_j - \sum_{j=\xi_{\ell_1}+1}^{i_0} p_j \right| \\ &\leq \left| \sum_{j=\xi_{\ell_1}+1}^{i_0} \bar{p}_j - \sum_{j=\xi_{\ell_1}+1}^{i_0} \hat{p}_j \right| + \left| \sum_{j=\xi_{\ell_1}+1}^{i_0} \hat{p}_j - \sum_{j=\xi_{\ell_1}+1}^{i_0} p_j \right| \\ &\leq \frac{\sum_{j=\xi_{\ell_1}+1}^{i_0} \hat{p}_j}{\sum_{j=\xi_{\ell_1}+1}^{\xi_{\ell_1+1}} \hat{p}_j} \left| \sum_{j=\xi_{\ell_1}+1}^{\xi_{\ell_1+1}} \hat{p}_j - \sum_{j=\xi_{\ell_1}+1}^{\xi_{\ell_1+1}} q_j \right| + \left| \sum_{j=\xi_{\ell_1}+1}^{i_0} \hat{p}_j - \sum_{j=\xi_{\ell_1}+1}^{i_0} p_j \right| \\ &\leq \left| \sum_{j=\xi_{\ell_1}+1}^{\xi_{\ell_1+1}} \hat{p}_j - \sum_{j=\xi_{\ell_1}+1}^{\xi_{\ell_1+1}} p_j \right| + \left| \sum_{j=\xi_{\ell_1}+1}^{i_0} \hat{p}_j - \sum_{j=\xi_{\ell_1}+1}^{i_0} p_j \right|. \end{aligned} \tag{A.4}$$

Note that ℓ_1 depends on m as well as i_0 . This means that level set containing i_0 may change as m varies. Now we consider all possible level sets containing $i = \eta_{\ell_0}$ for $m \geq m_0(\omega)$. Such level sets are of form

$$\{\eta_\alpha + 1, \dots, \eta_\beta\}, \quad \text{for } 0 \leq \alpha < \ell_0 \leq \beta \leq A.$$

Hence we have

$$\left| \sum_{j=\xi_{\ell_1}+1}^{\xi_{\ell_1+1}} \hat{p}_j - \sum_{j=\xi_{\ell_1}+1}^{\xi_{\ell_1+1}} p_j \right| \leq \max_{0 \leq \alpha < \ell_0 \leq \beta \leq A} \left| \sum_{j=\eta_\alpha+1}^{\eta_\beta} \hat{p}_j - \sum_{j=\eta_\alpha+1}^{\eta_\beta} p_j \right| \quad \text{and}$$

$$\left| \sum_{j=\xi_{\ell_1}+1}^{i_0} \hat{p}_j - \sum_{j=\xi_{\ell_1}+1}^{i_0} p_j \right| \leq \max_{0 \leq \alpha < \ell_0} \left| \sum_{j=\eta_\alpha+1}^{i_0} \hat{p}_j - \sum_{j=\eta_\alpha+1}^{i_0} p_j \right|.$$

Then (A.4) becomes

$$\left| \sum_{j=1}^{i_0} \bar{p}_j - \sum_{j=1}^{i_0} q_j \right| \leq \max_{0 \leq \alpha < \ell_0 \leq \beta \leq A} \left| \sum_{j=\eta_\alpha+1}^{\eta_\beta} \hat{p}_j - \sum_{j=\eta_\alpha+1}^{\eta_\beta} p_j \right| + \max_{0 \leq \alpha < \ell_0} \left| \sum_{j=\eta_\alpha+1}^{i_0} \hat{p}_j - \sum_{j=\eta_\alpha+1}^{i_0} p_j \right|.$$

This completes the proof. \square

References

- Billingsley, P. (1986). *Probability and Measure*, 2nd ed., Wiley, New York.
- Kiefer, J. (1961). On large deviation of the empirical distribution function of vector random variable and a law of the iterative logarithm, *Pacific Journal of Mathematics*, **11**, 649–660.
- Lee, C. I. C. (1987). Maximum Likelihood Estimates for Stochastically Ordered Multinomial Populations with Fixed and Random Zeros, *Foundation of Statistical Inference*, (I. B. MacNeill and G. J. Umphrey eds.) 189–197.
- Robertson, T. and Wright, F. T. (1974a). On the maximum likelihood estimation of stochastically ordered random variables, *The Annals of Statistics*, **2**, 528–534.
- Robertson, T. and Wright, F. T. (1974b). A norm reducing property for isotonized cauchy mean value functions, *The Annals of Statistics*, **2**, 1302–1307.
- 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.
- Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*, Wiley, Chichester.

Received February 16, 2013; Revised April 30, 2013; Accepted May 8, 2013