Statistica Sinica **14**(2004), 283-295

THE MAXIMUM LIKELIHOOD ESTIMATES OF EXPECTED FREQUENCIES UNDER THE LOOP ORDER

Ning-Zhong Shi and Shu-Rong Zheng

Northeast Normal University and Jilin University

Abstract: In this paper, we study the maximum likelihood estimates (MLEs) of expected frequencies under a loop order restriction in an $I \times J$ contingency table. Some properties of the MLEs are given and an algorithm for computing the MLEs is detailed. The proposed methods are illustrated by using the data from Wing (1962).

Key words and phrases: Expected frequency, Kuhn-Tucker condition, local odds ratio, loop order, MLE.

1. Introduction

Consider a 3×3 contingency table with observed frequencies m_{ij} and corresponding expected frequencies μ_{ij}, $i,\ j = 1, 2, 3$. The local odds ratios ψ_{ij} are written as $\psi_{ij} = (\mu_{ij} \times \mu_{i+1,j+1})/(\mu_{i,j+1} \times \mu_{i+1,j})$ for $i, j = 1, 2$. The local odds ratios are said to obey a loop order if the inequalities $\psi_{11} \leq \psi_{12} \leq \psi_{22}$ and $\psi_{11} \leq \psi_{21} \leq \psi_{22}$ are satisfied. We write this as the loop order by

$$\psi_{11} \leq \psi_{12}, \psi_{21} \leq \psi_{22}. \tag{1}$$

The problem is to find the MLE of μ_{ij} under the above restriction.

Fienberg (1978), Haberman (1974), Simon (1974), Goodman (1979), Agresti (1987) and others formulate associations based on scores of column and row effects in contingency tables and obtain estimates of expected frequencies. Loop order is a particular model in their papers. We consider estimating expected frequencies under the loop order without assigning the scores of row and column effects. Darroch and Ratcliff (1972) give an algorithm for generalized scaling loglinear models when the expected frequencies are restricted by some equalities. For the case of local homogeneous odds ratios several authors, such as Agresti (1984) and Yanagawa and Fujii (1990, 1995), consider estimating expected frequencies and give an algorithm to obtain MLEs. For $2 \times r$ contingency tables, Shi (1991) discusses MLEs when the simple order or loop order of odds ratios is satisfied. Agresti and Coull (1998) consider the order-restricted inference for monotone trend alternatives in contingency table. Lemke and Dykstra (1984) propose an algorithm of the multinomial maximum likelihood estimates with multiple cone restrictions. In our paper, based on iterative proportional fitting (Darroch and Ratcliff (1972)) and isotonic regression, an algorithm to obtain the estimates under the loop order is presented. Some results about isotonic regression are given in Section 2. Iterative proportional fitting (IPF) was first presented by Deming and Stephan (1940), further details on uses of IPF, see Agresti (1990, pp.185-186), (1984, pp.65-66) and Bishop, Fienberg and Holland (1975, pp.76-102).

In Section 2 of this paper, isotonic regression is reviewed. Section 3 describes the existence and uniqueness of the MLE of μ_{ij}, $i, j = 1, 2, 3$ under the loop order restriction. Section 4 is devoted to the proposed algorithm for computing the MLEs. In Section 5, we show how the algorithm may be generalized to $I \times J$ tables. In Section 6, an example of the use of the algorithm is presented.

2. Isotonic Regression

We review certain results about isotonic regression which may be found in Robertson, Wright and Dykstra (1988). In n-dimensional Eucliean space R^n, define an inner product with weight $\omega = (\omega_1, \ldots, \omega_n)$ and a norm as

$$(y, z)_\omega = \sum_{i=1}^{n} y_i z_i \omega_i \tag{2}$$

for any $y, z \in R^n$, where $\omega_i \geq 0$, $i = 1, \ldots, n$, and $\sum_{i=1}^{n} \omega_i = 1$. The vector $\theta = (\theta_1, \ldots, \theta_n)$ is said to satisfy the simple loop order if $\theta_1 \leq \theta_i \leq \theta_n$ for $i = 2, \ldots, n-1$. Let G denote the loop order cone $G = \{\theta\}$ when θ satisfies the loop order.

Definition 1. Let $x \in R^n$. Then $x^* = (x_1^*, \ldots, x_n^*)$ is called isotonic regression of x on G with weight vector ω if $x^* \in G$ and $(x-x^*, x-x^*)_\omega = \min_{\theta \in G}(x-\theta, x-\theta)_\omega$.

Theorem 2.1. *If x^* is isotonic regression of x on G with weight vector ω if $(x - x^*, x^*)_\omega = 0$ and $(x - x^*, y)_\omega \leq 0$, for any $y \in G$.*

Corollary 2.1. *For any real function g, we have $(x - x^*, g(x^*))_\omega = 0$, where $g(x^*) = (g(x_1^*), \ldots, g(x_n^*))$.*

The pool-adjacent-violater algorithm (PAVA) is commonly used to compute the isotonic regression x^* of x on G with weight ω. Let $M = (M_1, \ldots, M_n)$ be the subscript set of x after permutating $(x_2, \ldots, x_{n-1})$ in increasing order with $M_1 = 1$ and $M_n = n$. Let $z = (z_1, \ldots, z_n)$ where $z_i = x_{M_i}$ for all i. Let $\tilde{\omega} = (\tilde{\omega}_1, \ldots, \tilde{\omega}_n)$ be the weight where $\tilde{\omega}_i = \omega_{M_i}$. Let $G_0 = \{\theta\}$ where θ satisfies the simple order $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_n$. Let $A_V(B) = \sum_{i \in B} z_i \omega_i / (\sum_{i \in B} \omega_i)$, where B

is a subset of $i, \ldots, n$. To obtain the isotonic regression z^* of z to G_0 under the weight $\tilde{\omega}$, one uses the following.

PAVA

Step 1. if $z \in G_0$, then $z^* = z$.

Step 2. if there exists j that satisfies $z_j > z_{j+1}$, let $B = \{j, j+1\}$, $z_B = A_V(B)$, $\tilde{\omega}_B = \tilde{\omega}_j + \tilde{\omega}_{j+1}$, $\hat{z} = (z_1, \ldots, z_{j-1}, z_B, z_{j+2}, \ldots, z_n)$ and $\hat{\omega} = (\tilde{\omega}_1, \ldots, \tilde{\omega}_{j-1}, \tilde{\omega}_B, \tilde{\omega}_{j+1}, \ldots, \tilde{\omega}_n)$.

Step 3. Repeat Step 2 until the subscripts are partitioned into l blocks $B_1, \ldots,$ B_l satisfying $A_V(B_1) < \cdots < A_V(B_l)$, then $z_i^* = A_V(B_t)$, $i \in B_t$, $t = 1, \ldots, l$.

3. The Model and Some Properties of MLEs

We assume that the data are distributed as the multinomial distribution. The log-likelihood function is of the form $L(\mu) = \sum_{i=1}^3 \sum_{j=1}^3 m_{ij} \log \mu_{ij} + c$, where

$$\sum_{i=1}^3 \sum_{j=1}^3 \mu_{ij} = \sum_{i=1}^3 \sum_{j=1}^3 m_{ij} = m. \tag{3}$$

Here $\mu = (\mu_{ij}, i = 1,2,3; j = 1,2,3)$ and c is a constant which does not depend on the parameters. Now we suppose, in particular, that $\log \mu_{11} = \alpha_0 + \alpha_1^r + \alpha_1^c$, $\log \mu_{12} = \alpha_0 + \alpha_1^r + \alpha_2^c$, $\log \mu_{13} = \alpha_0 + \alpha_1^r + \alpha_3^c$, $\log \mu_{21} = \alpha_0 + \alpha_2^r + \alpha_1^c$, $\log \mu_{22} = \alpha_0 + \alpha_2^r + \alpha_2^c + \psi_1$, $\log \mu_{23} = \alpha_0 + \alpha_2^r + \alpha_3^c + \psi_1 + \psi_2$, $\log \mu_{31} = \alpha_0 + \alpha_3^r + \alpha_1^c$, $\log \mu_{32} = \alpha_0 + \alpha_3^r + \alpha_2^c + \psi_1 + \psi_3$, $\log \mu_{33} = \alpha_0 + \alpha_3^r + \alpha_3^c + \psi_1 + \psi_2 + \psi_3 + \psi_4$. Assume that $\sum_{i=1}^3 \alpha_i^r = \sum_{j=1}^3 \alpha_j^c = 0$, i.e., $\alpha_3^r = -\alpha_1^r - \alpha_2^r$ and $\alpha_3^c = -\alpha_1^c - \alpha_2^c$. The loop order in our particular case has

$$\psi_1 \le \psi_2, \psi_3 \le \psi_4. \tag{4}$$

For the convenience, the likelihood function $L(\mu)$ is denoted by $L(\alpha, \psi)$ where $\alpha = (\alpha_0, \alpha_1^r, \alpha_2^r, \alpha_1^c, \alpha_2^c)$ and $\psi = (\psi_1, \psi_2, \psi_3, \psi_4)$. It is clear that if (α^*, ψ^*) is the MLE of (α, ψ) in $L(\alpha, \psi)$ under (3) and (4), the corresponding values $\{\mu_{ij}^*\}$, based on the log-linear model, are the MLEs of $\{\mu_{ij}\}$ under (3) and (4). We call μ_{ij}^* the restricted MLE of μ_{ij}. Let

$$F(\alpha, \psi) = \sum_{i=1}^3 \sum_{j=1}^3 (m_{ij} \log \mu_{ij} - \mu_{ij}) + m. \tag{5}$$

The Hessian matrix of $F(\alpha, \psi)$ for α and ψ is negative definite, so $F(\alpha, \psi)$ is strictly concave for α and ψ. This implies that there exists a unique point (α^*, ψ^*) satisfying $F(\alpha^*, \psi^*) = \max_{\alpha,\psi} F(\alpha, \psi)$. In Appendix, a proof may be found to show that the maximum point (α^*, ψ^*) of $F(\alpha, \psi)$ under (4) is just the MLEs of (α, ψ) in $L(\alpha, \psi)$ under (3) and (4). We introduce the notation $A_1 = m_{22} + m_{23} + m_{32} + m_{33}$, $A_2 = m_{23} + m_{33}$, $A_3 = m_{32} + m_{33}$, $A_4 = m_{33}$. By substituting μ^*_{ij} for m_{ij} in A_i, we obtain A^*_i, i=1, 2, 3, 4. For instance, $A^*_1 = \mu^*_{22} + \mu^*_{23} + \mu^*_{32} + \mu^*_{33}$.

We need to find the solution (α^*, ψ^*) which maximizes $F(\alpha, \psi)$ under (3). The Lagrangian

$$\begin{aligned} -F(\alpha, \psi, \lambda) &= -F(\alpha, \psi) + \lambda_1(\psi_1 - \psi_2) + \lambda_2(\psi_1 - \psi_3) + \lambda_3(\psi_2 - \psi_4) + \lambda_4(\psi_3 - \psi_4) \\ &= -\sum_{i=1}^{3}\sum_{j=1}^{3}(m_{ij}\log\mu_{ij} - \mu_{ij}) - m + \lambda_1(\psi_1 - \psi_2) + \lambda_2(\psi_1 - \psi_3) \\ &\quad + \lambda_3(\psi_2 - \psi_4) + \lambda_4(\psi_3 - \psi_4), \end{aligned}$$

where $\lambda = (\lambda_1, \lambda_1, \lambda_3, \lambda_4)$ and the $\lambda_i's$ are the Lagrangian multipliers. The Kuhn-Tucker conditions, (see Mokhtar and Shetty (1979) or Anthony, Francis and Uhl Jr. (1992)) are usually used to deal with such problems. As $-F(\alpha, \psi)$ is a strictly convex function, (α^*, ψ^*) is the solution if

$$\begin{cases} \psi^*_1 \le \psi^*_2,\ \ \psi^*_3 \le \psi^*_4, \\ \lambda_i \ge 0, \quad i = 1, 2, 3, 4, \\ -\dfrac{\partial F}{\partial \alpha_0}\Big|_{(\alpha^*, \psi^*)} = -\sum_{i=1}^{3}\sum_{j=1}^{3}(m_{ij} - \mu^*_{ij}) = -m + \sum_{i=1}^{3}\sum_{j=1}^{3}\mu^*_{ij} = 0, \\ -\dfrac{\partial F}{\partial \alpha^r_i}\Big|_{(\alpha^*, \psi^*)} = -m_{i+} + \mu^*_{i+} + m_{3+} - \mu^*_{3+} = 0, \qquad i = 1, 2, \\ -\dfrac{\partial F}{\partial \alpha^c_j}\Big|_{(\alpha^*, \psi^*)} = -m_{+j} + \mu^*_{+j} + m_{+3} - \mu^*_{+3} = 0, \qquad j = 1, 2, \\ -\dfrac{\partial F}{\partial \psi_1}|_{(\alpha^*, \psi^*)} = -(m_{22} + m_{23} + m_{32} + m_{33} - \mu^*_{22} - \mu^*_{23} - \mu^*_{32} - \mu^*_{33}) + \lambda_1 + \lambda_2 = 0, \\ -\dfrac{\partial F}{\partial \psi_2}\Big|_{(\alpha^*, \psi^*)} = -(m_{23} + m_{33} - \mu^*_{23} - \mu^*_{33}) - \lambda_1 + \lambda_3 = 0, \\ -\dfrac{\partial F}{\partial \psi_3}\Big|_{(\alpha^*, \psi^*)} = -(m_{32} + m_{33} - \mu^*_{32} - \mu^*_{33}) - \lambda_2 + \lambda_4 = 0, \\ -\dfrac{\partial F}{\partial \psi_4}\Big|_{(\alpha^*, \psi^*)} = -(m_{33} - \mu^*_{33}) - \lambda_3 - \lambda_4 = 0, \\ \lambda_1(\psi_2 - \psi_1) = 0, \\ \lambda_2(\psi_3 - \psi_1) = 0, \\ \lambda_3(\psi_4 - \psi_2) = 0, \\ \lambda_4(\psi_4 - \psi_3) = 0. \end{cases}$$

The above equations correspond to

$(A)\ \psi_1^* \le \psi_2^*,\ \psi_3^* \le \psi_4^*,$

$$(B)\begin{cases} m_{i+} = \mu_{i+}^*, & i = 1,2,3 \\ m_{+j} = \mu_{+j}^*, & j = 1,2,3, \end{cases}$$

$$(C)\begin{cases} A_1 \ge A_1^* \quad (*1) \\ A_1 + A_2 \ge A_1^* + A_2^* \quad (*2) \\ A_1 + A_3 \ge A_1^* + A_2^* \quad (*3) \\ A_1 + A_2 + A_3 \ge A_1^* + A_2^* + A_3^* \quad (*4) \\ A_1 + A_2 + A_3 + A_4 = A_1^* + A_2^* + A_3^* + A_4^* \quad (*5) \\ " = "\text{holds in } \ (*1) \ \text{ if } \ \psi_1^* < \psi_2^* \ \text{ and } \ \psi_1^* < \psi_3^* \\ " = "\text{holds in } \ (*2) \ \text{ if } \ \psi_1^* < \psi_3^* \ \text{ and } \ \psi_2^* < \psi_4^* \\ " = "\text{holds in } \ (*3) \ \text{ if } \ \psi_1^* < \psi_2^* \ \text{ and } \ \psi_3^* < \psi_4^* \\ " = "\text{holds in } \ (*4) \ \text{ if } \ \psi_2^* < \psi_4^* \ \text{ and } \ \psi_3^* < \psi_4^*. \end{cases}$$

Finally, we have the following theorem.

Theorem 3.1. *(α^*, ψ^*) is the unique point satisfying $F(\alpha^*, \psi^*) = \max_{(\alpha,\psi)\in G} F(\alpha, \psi)$ if (A), (B) and (C) hold.*

4. The Iterative Algorithm

In this section, we propose an iterative algorithm based on Theorem 3.1 to compute the MLEs of μ_{ij} under the loop restriction. In step $(n, 1)$ and $(n, 2)$, similar to IPF, we mainly want to assure that the sum of expected frequencies in every row and column is that of observation frequencies. In Step $(n, 3)$, our main thought is to place weights on the μ_{ij} in order to assure the loop order. Let the starting point $\mu_{ij}^{(0,3)}$ be m_{ij} for $i = 1, 2, 3,\ j = 1, 2, 3$. As one cycle has three steps, we go to the stage where Step $(n-1, 1)$, Step $(n-1, 2)$ and Step $(n-1, 3)$ are satisfied.

Step$(n, 1)$ $\mu_{ij}^{(n,1)} = \mu_{ij}^{(n-1,3)} m_{i+}/\mu_{i+}^{(n-1,3)}$ for $i = 1, 2, 3,\ j = 1, 2, 3.$

Step$(n, 2)$ $\mu_{ij}^{(n,2)} = \mu_{ij}^{(n,1)} m_{+j}/\mu_{+j}^{(n,1)}$ for $i = 1, 2, 3,\ j = 1, 2, 3.$

$$\text{Step}(n,3)\begin{cases}\mu_{ij}^{(n,3)}=\mu_{ij}^{(n,2)}\times\dfrac{\hat{A}}{A^n},\text{for}\quad i=1\quad\text{or}\quad j=1,\\ \mu_{22}^{(n,3)}=\mu_{22}^{(n,2)}\times(\dfrac{a_1^{(n)}}{\psi_1^{(n-1,3)}})^t\times(\dfrac{\hat{A}}{A^n})^{1-t},\\ \mu_{23}^{(n,3)}=\mu_{23}^{(n,2)}\times(\dfrac{a_2^{(n)}}{\psi_2^{(n-1,3)}})^t\times(\dfrac{a_1^{(n)}}{\psi_1^{(n-1,3)}})^t\times(\dfrac{\hat{A}}{A^n})^{1-2t},\\ \mu_{32}^{(n,3)}=\mu_{32}^{(n,2)}\times(\dfrac{a_3^{(n)}}{\psi_3^{(n-1,3)}})^t\times(\dfrac{a_1^{(n)}}{\psi_1^{(n-1,3)}})^t\times(\dfrac{\hat{A}}{A^n})^{1-2t},\\ \mu_{33}^{(n,3)}=\mu_{33}^{(n,2)}\times(\dfrac{a_4^{(n)}}{\psi_4^{(n-1,3)}})^t\times(\dfrac{a_3^{(n)}}{\psi_3^{(n-1,3)}})^t\times(\dfrac{a_2^{(n)}}{\psi_2^{(n-1,3)}})^t\\ \qquad\times(\dfrac{a_1^{(n)}}{\psi_1^{(n-1,3)}})^t\times(\dfrac{\hat{A}}{A^n})^{1-4t}.\end{cases}$$

where $t=1/9$ and $\hat{A}=\sum_{j=1}^{3}m_{1j}+\sum_{i=2}^{3}m_{i1}+8/9m_{22}+7/9m_{23}+7/9m_{32}+5/9m_{33}$.

By substituting $\{\mu_{ij}^{(n,2)}\}$ for m_{ij} in $\hat{A}$ and A_i, we obtain A^n and $A_i^{(n)}$, by substituting $\{\mu_{ij}^{(n-1,3)}\}$ for μ_{ij} in ψ_{ij}, we obtain the corresponding $\psi_i^{(n-1,3)}$. For example, $\psi_1^{(n-1,3)}=\mu_{11}^{(n-1,3)}\times\mu_{22}^{(n-1,3)}/(\mu_{12}^{(n-1,3)}\times\mu_{21}^{(n-1,3)})$. The isotonic regression of $x^{(n)}=(x_i^{(n)},\ i=1,2,3,4)$ onto G with the weight vector $\omega^{(n)}=(\omega_i^{(n)},\ i=1,2,3,4)$ is denoted by $\vec{a}^{(n)}=P_{\omega^{(n)}}(x^{(n)}|G)$, where $\vec{a}^{(n)}=(a_i^{(n)},\ i=1,2,3,4)$ and G is the convex cone formed by the loop order restriction (4). Note that, in the above algorithm, the weights $\omega_i^{(n)}$ are $A_i^{(n)}$, $i=1,2,3,4$, and $x_i^{(n)}$ is denoted by $A_i\psi_i^{(n-1,3)}/A_i^{(n)}$, $i=1,2,3,4$. The $\mu_{ij}^{(n,1)}$ satisfy $\mu_{i+}^{(n,1)}=m_{i+}$ for all i, and the $\mu_{ij}^{(n,2)}$ satisfy $\mu_{+j}^{(n,1)}=m_{+j}$ for all j. Furthermore, we have $\psi_i^{(n,2)}=\psi_i^{(n,1)}=\psi_i^{(n-1,3)}$ and $\psi_i^{(n,3)}=(\psi_i^{(n-1,3)})^{8/9}(a_i^{(n)})^{1/9}(A^n/\hat{A})^{1/9}$, $i=1,2,3,4$.

Now, we explain Step$(n,3)$. We propose Step$(n,3)$ for Conditions (A) and (C). First, since $(a_i^{(n)},\ i=1,2,3,4)$ is the isotonic regression of $x^{(n)}$ onto G, we have $a_1^{(n)}\le a_2^{(n)}$, $a_3^{(n)}\le a_4^{(n)}$. Thus $\psi_i^{(n,3)}$ satisfies $\psi_1^{(n,3)}\le\psi_2^{(n,3)}$, $\psi_3^{(n,3)}\le\psi_4^{(n,3)}$, that is, the above algorithm assures the loop order restriction in each step. Furthermore, for Condition (C), we want to place weights on $\mu_{ij}^{(n,2)}$. Now μ_{22}^* only exists in A_1^*, μ_{23}^* exists in both A_1^* and A_2^*, μ_{23}^* exists in both A_1^* and A_3^*, and μ_{33}^* exists in all A_i^*, $i=1,2,3,4$, allowing Step$(n,3)$. The weight t in Step$(n,3)$, satisfies $0<4t<1$, the exponents of $a_i^{(n)}/\psi_i^{(n-1,3)}$ are all the same and the sum of exponents of weights in every $\mu_{ij}^{(n,3)}$ is 1. In our algorithm, without loss of generality, we set $t=1/9$.

Proposition 1. *If* $\sum_{i=1}^{k}p_i\ge\sum_{i=1}^{k}r_i$*, where* $p_i,r_i>0$*, for all* i*, then* $\sum_{i=1}^{k}p_i\log$

$(p_i/r_i) \geq 0$. *Furthermore,* $\sum_{i=1}^k p_i \log(p_i/r_i) = 0$ *if and only if* $p_i = r_i$, $i = 1,\ldots,k$ (*See Lemke and Dykstra* (1984)).

In order to guarantee that this algorithm sequence converges to the MLE, we need some lemmas. Proofs of Lemmas 1, 2 and 3 are given in the Appendix.

Lemma 1. *For any* $n \geq 1$, *we have* $\sum_{i=1}^3 \sum_{j=1}^3 \mu_{ij}^{(n,3)} \leq m$ *and* $\sum_{i=1}^3 \sum_{j=1}^3 \mu_{ij}^{(n,1)} = \sum_{i=1}^3 \sum_{j=1}^3 \mu_{ij}^{(n,2)} = m$.

Let $W_1^{(n)} = \sum_{i=1}^3 m_{i+} \log(m_{i+}/\mu_{i+}^{(n-1,3)})$, $W_2^{(n)} = \sum_{j=1}^3 m_{+j} \log(m_{+j}/\mu_{+j}^{(n-1,3)})$, $W_3^{(n)} = (1/9)\sum_{i=1}^4 A_i \log(a_i^{(n)}/\psi_i^{(n-1,3)}) + \hat{A}\log(\hat{A}/A^n)$ and $m^{(n,2)} = \sum_{i=1}^3 \sum_{j=1}^3 \mu_{ij}^{(n,2)} = m$.

Lemma 2. $W_i^{(n)} \geq 0$ *for* $i = 1, 2$ *or* 3.

Theorem 4.1. $\sum_{i=1}^3 \sum_{j=1}^3 m_{ij} \log \mu_{ij}^{(n,s)}$ *is increasing in* n *and converges for fixed* s, $s = 1, 2, 3$, *that is, the likelihood function converges.*

Proof. We have $W_1^{(n)} = \sum_{i=1}^3 \sum_{j=1}^3 m_{ij} \log(\mu_{ij}^{(n,1)}/\mu_{ij}^{(n-1,3)})$, $W_2^{(n)} = \sum_{i=1}^3 \sum_{j=1}^3 m_{ij} \log(\mu_{ij}^{(n,2)}/\mu_{ij}^{(n,1)})$ and $W_3^{(n)} = \sum_{i=1}^3 \sum_{j=1}^3 m_{ij} \log(\mu_{ij}^{(n,3)}/\mu_{ij}^{(n,2)})$. So by Lemma 2, $\sum_{i=1}^3 \sum_{j=1}^3 m_{ij} \log \mu_{ij}^{(n,s)}$ is increasing in n for any fixed s, where $s = 1, 2, 3$. By Lemma 1 and Proposition 1, we have $\sum_{i=1}^3 \sum_{j=1}^3 m_{ij} \log(m_{ij}/\mu_{ij}^{(n,s)}) \geq 0$. Then $\sum_{i=1}^3 \sum_{j=1}^3 m_{ij} \log(\mu_{ij}^{(n,s)})$ is bounded by $\sum_{i=1}^3 \sum_{j=1}^3 m_{ij} \log m_{ij}$ for $s = 1, 2, 3$. So $\sum_{i=1}^3 \sum_{j=1}^3 m_{ij} \log \mu_{ij}^{(n,s)}$ converges, that is, the likelihood function converges for any fixed s, where $s = 1, 2, 3$.

Lemma 3. *If all* m_{ij} $(i = 1, 2, 3,\ j = 1, 2, 3)$ *are positive, then* $\lim_{n\to+\infty} (a_i^{(n)}/\psi_i^{(n-1,3)}) = 1$, $i = 1, 2, 3$.

Theorem 4.2. *If all* m_{ij} *are positive, the sequence* $\{\mu_{ij}^{(n)}\}$ *given in the algorithm converges to the uniquely restricted MLEs of* $\{\mu_{ij}\}$.

Proof. Since $\mu_{ij}^{(n,s)}$ is bounded by Lemma 1, for any sequence there exists a subsequence, say $\{n_q\}$, which satisfies $\lim_{q\to+\infty} \mu_{ij}^{(n_q,s)} = \mu_{ij}^*$, $s = 1, 2, 3$). By the proposed algorithm, we have $\sum_{j=1}^3 \mu_{ij}^{(n_q,1)} = m_{i+}$. Hence $\mu_{i+}^* = m_{i+}$ and $\mu_{+j}^* = m_{+j}$. From $\psi_1^{(n_q,3)} \leq \psi_2^{(n_q,3)}, \psi_3^{(n_q,3)} \leq \psi_4^{(n_q,3)}$, one has $\psi_1^* \leq \psi_2^*, \psi_3^* \leq \psi_4^*$. This implies that μ_{ij}^* satisfies (A) and (B), and $\psi = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)$ falls under one of nine conditions:

$$\begin{array}{lll} \psi_1^* = \psi_2^* = \psi_3^* = \psi_4^*; & \psi_1^* = \psi_2^* < \psi_3^* < \psi_4^*; & \psi_1^* = \psi_3^* < \psi_2^* < \psi_4^*; \\ \psi_1^* < \psi_2^* = \psi_3^* < \psi_4^*; & \psi_1^* < \psi_2^* < \psi_3^* = \psi_4^*; & \psi_1^* < \psi_3^* < \psi_2^* = \psi_4^*; \\ \psi_1^* = \psi_2^* = \psi_3^* < \psi_4^*; & \psi_1^* < \psi_2^* = \psi_3^* = \psi_4^*; & \psi_1^* < \psi_2^* < \psi_3^* < \psi_4^*. \end{array}$$

For the situation $\psi_1^* = \psi_2^* < \psi_3^* < \psi_4^*$, we prove that $\{\mu_{ij}^*\}$ satisfies (C). Since $\psi_i^{(n_q,3)} = (\psi_i^{(n_q-1,3)})^{8/9}(a_i^{(n_q)})^{1/9}(A^{n_q}/\hat{A})^{1/9}$, $i=1,2,3$, then by Lemma 3, $\lim_{q\to+\infty} \psi_i^{(n_q-1,3)} = \lim_{q\to+\infty} \psi_i^{(n_q,3)} = \psi_i^* = \lim_{q\to+\infty} a_i^{(n_q)}$. So there exists a sufficiently large q satisfying $a_1^{(n_q)} \le a_2^{(n_q)} < a_3^{(n_q)} < a_4^{(n_q)}$. By the PAVA algorithm of isotonic regression (see Robertson, Wright and Dykstra (1988)), we obtain $a_3^{(n_q)} = x_3^{(n_q)}$, $a_4^{(n_q)} = x_4^{(n_q)}$, that is, $A_i\psi_i^{(n_q-1,3)}/a_i^{(n_q)} = A_i^{(n_q)}$, $i=3,4$. By Corollary 2.1 where $g(x*) = (1.0/a_1^{(n_q)}, \ldots, 1.0/a_4^{(n_q)})$, we have $\sum_{i=1}^4 A_i\psi_i^{(n_q-1,3)}/a_i^{(n_q)} = \sum_{i=1}^4 A_i^{(n_q)}$. Then $A_1\psi_1^{(n_q-1,3)}/a_1^{(n_q)} + A_2\psi_2^{(n_q-1,3)}/a_2^{(n_q)} = A_1^{(n_q)} + A_2^{(n_q)}$ and $x_1^{(n_q)} \ge a_1^{(n_q)}$, that is, $A_1^{(n_q)} a_1^{(n_q)} \le A_1\psi_1^{(n_q-1)}$. Furthermore, we have

$$\begin{aligned} A_1 + A_2 - (A_1^* + A_2^*) &= \lim_{q\to+\infty}\Big(\frac{A_1\psi_1^{(n_q-1,3)}}{a_1^{(n_q)}} + \frac{A_2\psi_2^{(n_q-1,3)}}{a_2^{(n_q)}}\Big) - (A_1^* + A_2^*) \\ &= \lim_{q\to+\infty}(A_1^{(n_q)} + A_2^{(n_q)}) - (A_1^* + A_2^*) = 0, \end{aligned}$$

$$\begin{aligned} A_1 - A_1^* = A_1 - \lim_{q\to+\infty} A_1^{(n_q)}\frac{a_1^{(n_q)}}{\psi_1^{(n_q)}} &= A_1 - \lim_{q\to+\infty} A_1^{(n_q)} \times \frac{a_1^{(n_q)}}{\psi_1^*} \\ \ge A_1 - \lim_{q\to+\infty} A_1\frac{\psi_1^{(n_q-1,3)}}{\psi_1^*} &= A_1 - A_1 = 0. \end{aligned}$$

By Lemma 3, we have $A_4 - A_4^* = A_4 - \lim_{q\to+\infty} A_4^{(n_q)} = A_4 - \lim_{q\to+\infty} A_4 \times (\psi_4^{(n_q-1,3)}/a_4^{(n_q)}) = A_4 - A_4 = 0$. Similarly $A_3 - A_3^* = 0$. So we have $A_1 \ge A_1^*$, $A_1 + A_2 = A_1^* + A_2^*$, $A_1 + A_3 \ge A_1^* + A_2^*$, $A_1 + A_2 + A_3 = A_1^* + A_2^* + A_3^*$, $A_1 + A_2 + A_3 + A_4 = A_1^* + A_2^* + A_3^* + A_4^*$.

The remaining eight situations are proved by the same method. So $\{\mu_{ij}^*\}$ satisfy (C) of Theorem 3.1, that is, $\{\mu_{ij}^*\}$ are the MLE of $\{\mu_{ij}\}$ under (4).

5. A Generalization of the Algorithm for an $I \times J$ Table

Now we show that the algorithm and the results in Section 4 are also suitable for an $I \times J$ table. Here the loop order is

$$\psi_{ij} \le \psi_{kl} \qquad \text{for} \qquad i + j < k + l. \tag{6}$$

The following algorithm will provide us with a useful tool for computing the MLE of the expected frequencies under the loop order. Fortunately, the theorem for convergence in Section 4 is also suitable for the present algorithm. Notations are similar to those in Section 4.

Start the algorithm with $\mu_{ij}^{(0,3)} = m_{ij}$ for all i and j.

Step$(n,1)$ $\mu_{ij}^{(n,1)} = \mu_{ij}^{(n-1,3)} m_{i+}/\mu_{i+}^{(n-1,3)}$, for all i and j.

Step$(n,2)$ $\mu_{ij}^{(n,2)} = \mu_{ij}^{(n,1)} m_{+j}/\mu_{+j}^{(n,1)}$, for all i and j.

$$\text{Step}(n,3) \begin{cases} \mu_{ij}^{(n,3)} = \mu_{ij}^{(n,2)} \hat{A}/A^n, \text{ for } \ i=1 \ \text{ or } \ j=1 \\ \mu_{ij}^{(n,3)} = \mu_{ij}^{(n,2)} \prod_{p,q=1}^{i,j} (\dfrac{a_{pq}^{(n)}}{\psi_{pq}^{(n-1,3)}})^{1/K} (\dfrac{\hat{A}}{A^n})^{(K-(i-1)(j-1))/K}, \text{ otherwise,} \end{cases}$$

where $\hat{A} = \sum_{i=2}^{I} m_{i1} + \sum_{j=2}^{J} m_{1j} + m_{11} + \sum_{i=2}^{I} \sum_{j=2}^{J} (K-(i-1)(j-1)) m_{ij}/K$ and $K = I(I-1)J(J-1)/4$. If we substitute $\mu_{ij}^{(n,2)}$ for m_{ij} in $\hat{A}$, we obtain $A^{(n)}$. Let $\vec{a}^{(n)} = (a_{ij}^{(n)}, i = 1, \ldots, I-1, j = 1, \ldots, J-1)$ denote the weighted isotonic regression of $x^{(n)} = (x_{ij}^{(n)})$ under G, where the weight vector is $\omega^{(n)} = (\omega_{ij}^{(n)})$ and G denotes the cone consisting of vectors restricted by the loop order(6). Here we let $\omega_{ij}^{(n)} = A_{ij}^{(n)}$, and $x_{ij}^{(n)}$ is denoted by $A_{ij}\psi_{ij}^{(n-1,3)}/A_{ij}^{(n)}$ where $A_{ij} = \sum_{p=i+1}^{I} \sum_{q=j+1}^{J} m_{pq}$ and $A_{ij}^{(n)} = \sum_{p=i+1}^{I} \sum_{q=j+1}^{J} \mu_{pq}^{(n,2)}$, for $i = 1, \ldots I-1$, $j = 1, \ldots J-1$.

Theorem 5.1. *If all m_{ij} are positive, the sequence $\{\mu_{ij}^{(n)}\}$ converges to the MLEs of $\{\mu_{ij}\}$ under the loop order* (6).

It is readily seen that the proof of Theorem 5.1 is the same as that of Theorem 4.2.

6. Numerical Example

For illustration, the proposed algorithm is used to study the data given by Wing (1962) comparing frequency of visits with length of stay for 132 long-term schizophrenic patients in two London mental hospitals (Table 1).

Table 1. Frequency of visits by length of stay for 132 long-term schizophrenic patients (Wing (1962)).

Frequency of visits	Length of stay in hospital			
	At least 20 years	at least 10 years but less than 20 years	at least 2 years but less than 10 years	Totals
Goes home, or visited regularly	3	16	43	62
Visited less than once a month. Does not go home	10	11	6	27
Never visited and never goes home	16	18	9	43
Totals	29	45	58	132

These data were analyzed by Haberman (1974), see also Fienberg (1978). From their models, we see that the odds ratios of the data satisfied $\psi_1 = \psi_2 < \psi_3 = \psi_4$, a special case of the loop order. In their papers, they first assigned the scores of row and column effects and then obtained the MLE of the expected frequencies. In this paper, we use a method different from that of Haberman (1974). If we only assume the loop order, what can we obtain the MLE of expected frequencies? We can obtain the MLE by the proposed algorithm. The computed results of estimating the expected frequencies are listed in Table 2.

Table 2. MLE of expected frequencies under the loop order for data in Table 1.

Frequency of visits	Length of stay in hospital		
	at least 20 years	at least 10 years but less than 20 years	at least 2 years but less than 10 years
Goes home, or visited regularly	2.93	16.22	42.85
Visited less than once a month. Does not go home	9.9	11.12	5.95
Never visited and never goes home	16.15	17.64	9.19

Acknowledgements

The authors thank the Co-editor, an associate editor and the referee for the helpful comments on the earlier version of this paper. We also thank Professors Bai Zhidong, Wang Yougan and Guo Jianhua for advice. The work is supported by the National Science Foundation of China.

Appendix

Remark. The maximum point of $F(\alpha, \psi)$ under (3) is just the MLE of (α, ψ) in $L(\alpha, \psi)$ under (2) and (3), that is, the maximum point of $F(\alpha, \psi)$ under (3) satisfies (2).

Proof. Let (α^*, ψ^*) be the maximum point of $F(\alpha, \psi)$ under (3), μ^*_{ij} is the corresponding expected frequency, $\lambda = m/\sum_{i=1}^3 \sum_{j=1}^3 \mu^*_{ij}$ and $\mu_{ij} = \lambda\mu^*_{ij}$. Then

$$\sum_{i=1}^3 \sum_{j=1}^3 \{m_{ij} \log \mu^*_{ij} - \mu^*_{ij}\} - \sum_{i=1}^3 \sum_{j=1}^3 \{m_{ij} \log \mu_{ij} - \mu_{ij}\}$$
$$= \lambda \sum_{i=1}^3 \sum_{j=1}^3 \mu^*_{ij} - \sum_{i=1}^3 \sum_{j=1}^3 \mu^*_{ij} - m \log \lambda$$

$$= (\lambda - 1)\sum_{i=1}^{3}\sum_{j=1}^{3}\mu_{ij}^{*} - m\log\lambda$$
$$= \lambda\sum_{i=1}^{3}\sum_{j=1}^{3}\mu_{ij}^{*}\{1 - 1/\lambda - \log\lambda\}.$$

Let $f(\lambda) = \lambda(1-1/\lambda-\log\lambda)$, so $f(\lambda) \le 0$ and $f(1) = 0$ if and only if $\lambda = 1$. Then the maximum point (α^*, ψ^*) of $F(\alpha, \psi)$ under (3) satisfies $\sum_{i=1}^{3}\sum_{j=1}^{3}\mu_{ij}^{*} = m$, that is, it satisfies (2).

Proof of Lemma 1. By Jensen's inequality, $x_1^{a_1} \cdot x_2^{a_2} \cdots x_n^{a_n} \le a_1x_1 + a_2x_2 + \cdots + a_nx_n$, where $x_i \ge 0$, $a_i \ge 0$, $i = 1, \ldots, n$ and $\sum_{i=1}^{n} a_i = 1$. Thus

$$\mu_{22}^{(n,3)} \le \frac{1}{9}\mu_{22}^{(n,2)}\frac{a_1^{(n)}}{\psi_1^{(n-1,3)}} + \frac{8}{9}\mu_{22}^{(n,2)}\frac{\hat{A}}{A^n}.$$
$$\mu_{23}^{(n,3)} \le \frac{1}{9}\mu_{23}^{(n,2)}\frac{a_2^{(n)}}{\psi_2^{(n-1,3)}} + \frac{1}{9}\mu_{23}^{(n,2)}\frac{a_1^{(n)}}{\psi_1^{(n-1,3)}} + \frac{7}{9}\mu_{23}^{(n,2)}\frac{\hat{A}}{A^n}.$$
$$\mu_{32}^{(n,3)} \le \frac{1}{9}\mu_{32}^{(n,2)}\frac{a_3^{(n)}}{\psi_3^{(n-1,3)}} + \frac{1}{9}\mu_{32}^{(n,2)}\frac{a_1^{(n)}}{\psi_1^{(n-1,3)}} + \frac{7}{9}\mu_{32}^{(n,2)}\frac{\hat{A}}{A^n}$$
$$\mu_{33}^{(n,3)} \le \frac{1}{9}\mu_{33}^{(n,2)}\frac{a_4^{(n)}}{\psi_4^{(n-1,3)}} + \frac{1}{9}\mu_{33}^{(n,2)}\frac{a_3^{(n)}}{\psi_3^{(n-1,3)}} + \frac{1}{9}\mu_{33}^{(n,2)}\frac{a_2^{(n)}}{\psi_2^{(n-1,3)}}$$
$$+\frac{1}{9}\mu_{33}^{(n,2)}\frac{a_1^{(n)}}{\psi_1^{(n-1,3)}} + \frac{5}{9}\mu_{33}^{(n,2)}\frac{\hat{A}}{A^n}.$$

By Theorem 2.1, $(x - x^*, y)_\omega \le 0$. Thus $\sum_{i=1}^{4}(A_i^{(n)}a_i^{(n)}/\psi_i^{(n-1,3)}) \le \sum_{i=1}^{4} A_i$ where $\omega = (A_1^{(n)}, A_2^{(n)}, A_3^{(n)}, A_4^{(n)})$, $x^* = (a_1^{(n)}, a_2^{(n)}, a_3^{(n)}, a_4^{(n)})$, $x = (x_1, x_2, x_3, x_4)$, where $x_i = A_iA_i^{(n)}/\psi_i^{(n-1,3)}$, and $y = (y_1, y_2, y_3, y_4)$ where $y_i = -1/\psi_i^{(n-1,3)}$. So we have

$$\mu_{22}^{(n,3)} + \mu_{23}^{(n,3)} + \mu_{32}^{(n,3)} + \mu_{33}^{(n,3)}$$
$$\le \frac{1}{9}\sum_{i=1}^{4}\frac{A_i^{(n)}a_i^{(n)}}{\psi_i^{(n-1,3)}} + \frac{\hat{A}}{A^n}(\frac{8}{9}\mu_{22}^{(n,2)} + \frac{7}{9}\mu_{23}^{(n,2)} + \frac{7}{9}\mu_{32}^{(n,2)} + \frac{5}{9}\mu_{33}^{(n,2)})$$
$$\le \frac{1}{9}\sum_{i=1}^{4}A_i + \frac{\hat{A}}{A^n}(\frac{8}{9}\mu_{22}^{(n,2)} + \frac{7}{9}\mu_{23}^{(n,2)} + \frac{7}{9}\mu_{32}^{(n,2)} + \frac{5}{9}\mu_{33}^{(n,2)}).$$

Hence $\sum_{i=1}^{3}\sum_{j=1}^{3}\mu_{ij}^{(n,3)} \le (1/9)\sum_{i=1}^{4} A_i + (\hat{A}/A^n)A^n = (1/9)\sum_{i=1}^{4} A_i + \hat{A} = m.$

Proof of Lemma 2. By Corollary 2.1, we have $\sum_{i=1}^4 (A_i\psi_i^{(n-1,3)}/a_i^{(n)}) = \sum_{i=1}^4 A_i^{(n)} = m^{(n,2)} - A^n$, where $\omega = (A_1^{(n)}, A_2^{(n)}, A_3^{(n)}, A_4^{(n)})$, $x = ((A_1/A_1^{(n)})\psi_1^{(n-1,3)}, \ldots, (A_4/A_4^{(n)})\psi_4^{(n-1,3)})$ and $x_* = (a_1^{(n)}, \ldots, a_4^{(n)})$, $y = (-1/a_1^{(n)}, \ldots, -1/a_4^{(n)})$. Since $-\log x$ is a convex function, we have by Jensen's inequality

$$\begin{aligned} W_3^{(n)} &= \frac{1}{9}\sum_{i=1}^4 A_i \log \frac{a_i^{(n)}}{\psi_i^{(n-1,3)}} + \hat{A}\log\frac{\hat{A}}{A^n} \\ &\geq -(m-\hat{A})\log\sum_{i=1}^4 \frac{\frac{1}{9}A_i\psi_i^{(n-1,3)}}{a_i^{(n)}(m-\hat{A})} + \hat{A}\log\frac{\hat{A}}{A^n} \\ &= (m-\hat{A})\log\frac{(m-\hat{A})}{m^{(n,2)}-A^n} + \hat{A}\log\frac{\hat{A}}{A^n}, \end{aligned}$$

where $m^{(n,2)} = m$. By Proposition 1, we have $W_i^{(n)} \geq 0$, $i = 1,2,3$.

Proof of lemma 3. By the expressions given for $W_i^{(n)}$, we have $\sum_{k=1}^3 W_k^{(n)} = \sum_{i=1}^{(3)}\sum_{i=1}^{(3)} m_{ij}\log(\mu_{ij}^{(n,3)}/\mu_{ij}^{(n-1,3)})$. By Theorem 1, $\lim_{n\to+\infty}\sum_{k=1}^3 W_k^{(n)} = 0$. Because $W_k^{(n)}$ is nonnegative, we have $\lim_{n\to+\infty} W_k^{(n)} = 0$ $k = 1,2,3$. From Proposition 1, we have $\lim_{n\to+\infty}\sum_{i=1}^4 A_i^{(n)} = \sum_{i=1}^4 A_i \lim_{n\to+\infty} A^n = \hat{A}$. Furthermore, because $\sum_{i=1}^4 A_i\psi_i^{(n-1,3)}/a_i^{(n)} = \sum_{i=1}^4 A_i^{(n)}$, we may have : $\exists\, M > 0$, $\forall n$, $\psi_i^{(n-1,3)}/a_i^{(n)} \leq M (i = 1,2,3,4)$. Then for any sequence, there exist a subsequence $\{n_j\}$ and a constant $a_i (i = 1,2,3,4)$ satisfying $\lim_{j\to+\infty}\psi_i^{(n_j-1,3)}/a_i^{(n_j)} = a_i$; $\lim_{n\to+\infty} W_3^{(n_j)} = 1/9\sum_{i=1}^4 A_i\log a_i^{-1} = 0$ and $\sum_{i=1}^4 A_i a_i = \lim_{j\to+\infty}\sum_{i=1}^4 A_i\psi_i^{(n_j-1,3)}/a_i^{(n_j)} = \lim_{j\to+\infty}\sum_{i=1}^4 A_i^{(n_j)} = \sum_{i=1}^4 A_i$. Since $-\log x$ is a strictly convex function, by Jensen's inequality we have

$$\begin{aligned} 0 &= \sum_{i=1}^4 A_i \log a_i^{-1} \\ &\geq -\sum_{i=1}^4 A_i\log(\sum_{i=1}^4 A_i a_i/\sum_{i=1}^4 A_i) \\ &= -\sum_{i=1}^4 A_i\log(\sum_{i=1}^4 A_i/\sum_{i=1}^4 A_i) = 0. \end{aligned}$$

Furthermore we have $a_1 = a_2 = a_3 = a_4 = 1$, so $\lim_{n\to+\infty}\psi_i^{(n-1,3)}/a_i^{(n)} = 1$, $i = 1,2,3$.

References

Agresti, A. (1984). *Analysis of Ordinal Categorical Data*. John Wiley and Sons, New York.

Agresti, A. (1990). *Categorical Data Analysis*. John Wiley and Sons, New York.

Agresti, A. and Coull, B. A. (2002). The analysis of contingency tables under inequality constraints. *J. Statist. Plann. Inference* **107**, 45-73.

Agresti, A., Chuang, C. and Kezouh, A. (1987). Order-restricted score parameters in association models for contingency tables. *J. Amer. Statist. Assoc.* **82**, 619-623.

Agresti, A. and Coull, B. A. (1998). Order-restricted inference for monotone trend alternatives in contingency table. *Comput. Statist. Data Anal.* **28**, 139-155.

Anthony, L. P., Francis, E. S. and Uhl Jr., J. J. (1992). *The Mathematics of Nonlinear Programming*. Springer-Verlag, New York.

Bishop, Y. M. M., Fienberg, S. E. and Holland, P. W. (1975). *Discrete Multivariate Analysis*. MIT Press, Cambridge, Massachusetts.

Darroch, J. N. and Ratcliff, D. (1972). Generalized iterative scaling for log-linear models. *Ann. Statist.* **43**, 1470-1480.

Deming, W. E. and Stephan, F. F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Statist.* **11**, 427-444.

Fienberg, S. E. (1978). *The Analysis of Cross-Classified Categorical Data*. MIT Press, Cambridge, Massachusetts.

Goodman, L. A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *J. Amer. Statist. Assoc.* **74**, 537-552.

Haberman, S. J. (1974). Loglinear models for frequency tables with ordered classifications. *Biometrics* **30**, 589-600.

Lemke, J. H. and Dykstra, R. L. (1984). An algorithm for multinormial maximum likelihood estimation with multiple cone restrictions. Techical Report 84-1, Department of Preventive Medicine, University of Iowa.

Mokhtar, S. B. and Shetty, C. M. (1979). *Nonlinear Programming: Theory and Algorithms*. John Wiley and Sons, New York.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Ordered Restricted Statistical Inference*. John Wiley and Sons, New York.

Shi, N.-Z.(1991). A test of homogeneity of odds ratios against order restrictions. *J. Amer. Statist. Assoc.* **86**, 154-158.

Simon, G. (1974). Alternative analysis for the singly-ordered contingency table. *J. Amer. Statist. Assoc.* **69**, 971-976.

Wing, J. K. (1962). Institutionalism in mental hospitals. *Brit. J. Soc. Clin. Psychol.* **1**, 38-51.

Meng, Xiao-Li and Dyk, David Van (1998). Fast EM-type implementations for mixed effects models. *J. Roy. Statist. Soc. Ser. B* **60**, 559-578.

Yanagaua, T. and Fujii, Y. (1990). Homogeneity test with a generalized Mantel-Haenszed estimator for L $2 \times J$ contingency table. *J. Amer. Statist. Assoc.* **85**, 744-748.

Yanagaua, T. and Fujii, Y. (1995). Projection-method Mantel-Haenszel estimator for k $2 \times J$ tables. *J. Amer. Statist. Assoc.* **90**, 649-656.

Department of Mathematics, Northeast Normal University, Changchun 130024, China.

E-mail: shinz@nenu.edu.cn

Department of Mathematics, Northeast Normal University, Changchun 130024, China.

E-mail: zhengsr@nenu.edu.cn

(Received September 2001; accepted February 2003)