Markov chain Monte Carlo exact tests for incomplete two-way contingency tables

Satoshi AOKI and Akimichi TAKEMURA

Graduate School of Information Science and Technology
University of Tokyo, Tokyo, Japan

SUMMARY

We consider testing the quasi-independence hypothesis for two-way contingency tables which contain some structural zero cells. For sparse contingency tables where the large sample approximation is not adequate, the Markov chain Monte Carlo exact tests are powerful tools. To construct a connected chain over the two-way contingency tables with fixed sufficient statistics and an arbitrary configuration of structural zero cells, an algebraic algorithm proposed by Diaconis and Sturmfels (1998) can be used. However their algorithm does not seem to be a satisfactory answer, because the Markov basis produced by the algorithm often contains many redundant elements and is hard to interpret. We derive an explicit characterization of a minimal Markov basis, prove its uniqueness and present an algorithm for obtaining the unique minimal basis. A computational example and the discussion on further basis reduction for the case of positive sufficient statistics are also given.

Keywords: Bradley-Terry model, Gröbner basis, minimal Markov basis, quasi-independence, quasi-symmetry, structural zero.

1 Introduction

Researchers often encounter the problem of analyzing incomplete contingency tables, i.e., tables containing some structural, or *a priori*, zeros. Structural zeros arise in situations where it is theoretically impossible for some cells to contain observations. For example, when a secondary infection can occur only if a primary infection occurs, the cell in the contingency table that corresponds to the secondary infection without the primary infection would necessarily contain a structural zero. Such cells can occur naturally as a feature of the data and can be distinguished from sampling zeros, which occur due to the sampling variability and the relative smallness of the cell probabilities. It is noted that if probabilities of some cells are free from models, i.e., if the probabilities of some cells are regarded as nuisance parameters, these cells can also be treated as if they are structural zero cells.

For analyzing two-way incomplete contingency tables, one of the most familiar models is the quasi-independence model (see Bishop *et al.*, 1975, for example). To perform the exact test of quasi-independence, the null distribution of an appropriate test statistic is required. However, a complete enumeration of this distribution is often computationally infeasible and a Monte Carlo exact test is used. See Agresti (1992) for example.

As another approach, a Markov chain Monte Carlo approach is extensively used in various settings of contingency tables, for example, Besag and Clifford (1989) for performing significance tests for the Ising model (two-way binary tables); Smith *et al.* (1996) for tests of independence, quasi-independence and quasi-symmetry for square contingency tables; Guo and Thompson (1992) for exact tests of Hardy-Weinberg proportions (triangular two-way contingency tables); Diaconis and Saloff-Coste (1995) for two-way contingency tables; Hernek (1998), Dyer and Greenhill (2000) for $2 \times J$ contingency tables; Forster *et al.* (1996) for 2^d contingency tables; Aoki and Takemura (2002) for $3 \times 3 \times K$ contingency tables; Diaconis and Sturmfels (1998) for general discrete exponential family. In this paper, we present Markov chain Monte Carlo algorithms for testing quasi-independence of incomplete two-way contingency tables. In addition we present similar results for testing quasi-symmetry for square tables.

Our main contribution is twofold. First, our work is an extension of the work of Smith *et al.* (1996), which considers the situation that the contingency table is square and the diagonal cells are structural zeros. In this paper, we consider more general situation; the contingency table is not necessarily square and there is no constraint on the configuration of structural zero cells. Second, our approach provides more concise and explicit expressions of a Markov basis than the general algorithms by Diaconis and Sturmfels (1998). Diaconis and Sturmfels (1998) proposed algebraic algorithms for finding a basis to construct a connected Markov chain over general multiway contingency tables with fixed sufficient statistic. Their algorithms involve computations in polynomial rings using Gröbner bases. However, algorithms based on Gröbner bases are very time consuming when the size of contingency tables is large. Furthermore, Gröbner basis computation produces large number of redundant basis elements due to the lack of symmetry and minimality inherent in Gröbner basis. Concerning the minimality of a basis, Takemura and Aoki (2002) gives some characterizations of a minimal basis and its uniqueness for a connected Markov chain for sampling from the conditional distribution. In this paper, we give the closed form expression of the unique minimal basis for two-way contingency tables with arbitrary configuration of structural zeros.

The construction of this paper is as follows. Section 2 and Section 3 describe the problem. Section 4 gives an explicit characterization a minimal basis. We also prove its uniqueness. Section 5 describes the algorithms for enumerating elements of the unique minimal basis. In this section, an explicit forms of minimal bases for some typical situations are also given. Computational example is given in Section 6. In Section 7 we discuss further basis reduction for the case of positive sufficient statistics. Finally in Appendix we give a corresponding result on quasi-symmetry for square two-way tables.

2 Exact tests for quasi-independence

Let $\mathbb{N} = \{0, 1, 2, \ldots\}$ and let $\boldsymbol{X}$ be an $I \times J$ contingency table with entries $x_{ij} \in \mathbb{N},\ i = 1, \ldots, I,\ j = 1, \ldots, J$. Let $S \subset \{(i, j) \mid 1 \leq i \leq I, 1 \leq j \leq J\}$ be the set of cells that are not structural zeros. We consider models for the cell probability in incomplete contingency tables in the (natural) logarithmic scale as

$$\log p_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} \qquad (1)$$

for $(i,j) \in S$ and $p_{ij} \equiv 0$ for $(i,j) \notin S$. We then define the model of quasi-independence for the subtable S by setting

$$\mathrm{H}_0 : \gamma_{ij} = 0$$

for $(i,j) \in S$. This is a natural extension of the familiar model of independence of the variables corresponding to rows and columns in the ordinary two-way contingency tables. An interpretation of this model and restrictions on the parameters are discussed in detail in Bishop *et al.* (1975).

For testing a null hypothesis in the presence of nuisance parameters, a common approach is to base the inference on the conditional distribution given a sufficient statistic for the nuisance parameters. Using this conditional distribution, an exact test can be constructed. See Agresti (1992) for details of exact tests for contingency tables. Under the null hypothesis of quasi-independence, the row sums, $x_{i\cdot}$, and the column sums, $x_{\cdot j}$, are the sufficient statistics for the nuisance parameters μ, α_i, β_j. The conditional distribution is then written as

$$f(\boldsymbol{X} \mid \{x_{i\cdot}\}, \{x_{\cdot j}\}, S) = C \prod_{(i,j)\in S} \frac{1}{x_{ij}!}, \tag{2}$$

where C is the normalizing constant determined from $\{x_{i\cdot}\}, \{x_{\cdot j}\}, S$ and written as

$$C^{-1} = \sum_{\boldsymbol{X} \in \mathcal{F}(\{x_{i\cdot}\},\{x_{\cdot j}\},S)} \left(\prod_{(i,j)\in S} \frac{1}{x_{ij}!} \right),$$

where

$$\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S) = \left\{ \boldsymbol{Y} \;\middle|\; \sum_{j=1}^{J} y_{ij} = x_{i\cdot}, \sum_{i=1}^{I} y_{ij} = x_{\cdot j},\; y_{ij} \in \mathbb{N},\; \text{and } y_{ij} = 0 \text{ for } (i,j) \notin S \right\}.$$

Hereafter, for simplicity we omit S in $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$, when $S = \{1, \ldots, I\} \times \{1, \ldots, J\}$, i.e., there is no structural zero cell. The discrepancy from the null hypothesis H_0 is measured by an appropriate test statistic. For each element in $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$, the value of this test statistic is calculated. The exact conditional p value is the sum of the conditional probabilities for the elements in $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$ which are at least as discrepant from the null hypothesis as the observed data.

Unfortunately, however, a complete enumeration of all the elements in $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$, and hence the calculation of the normalizing constant C, is usually computationally infeasible for large sparse tables. In this paper, instead, we consider a Markov chain Monte Carlo method described in the next section.

3 Metropolis-Hastings sampling

To perform the exact tests of quasi-independence, our approach is to generate samples from $f(\boldsymbol{X} \mid \{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$ and calculate the null distribution of various test statistics. If a connected Markov chain over $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$ is constructed, the chain can be modified to give a connected and aperiodic Markov chain with stationary distribution $f(\boldsymbol{X} \mid \{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$ by the usual Metropolis procedure (Hastings, 1970, for example).

If there is no structural zero cell, a connected Markov chain over $\mathcal{F}(\{x_{i\cdot}\},\{x_{\cdot j}\})$ is easily constructed as follows. Let $\boldsymbol{X}$ be the current state in $\mathcal{F}(\{x_{i\cdot}\},\{x_{\cdot j}\})$. The next state is selected by choosing a pair of rows and a pair of columns at random, and modifying $\boldsymbol{X}$ at the four entries where the selected rows and columns intersect as

$$\begin{array}{cc} + & - \\ - & + \end{array} \quad \text{or} \quad \begin{array}{cc} - & + \\ + & - \end{array} \quad \text{with probability } \frac{1}{2} \text{ each.} \tag{3}$$

The modification adds or subtracts 1 from each of the four entries, keeping the row and column sums. If the modification forces negative entries, discard it and continue by choosing a new pairs of rows and columns.

On the other hand, if there are structural zero cells, a chain constructed from this type of moves might not be connected. As the simplest example, consider 3×3 contingency tables having structural zero cells as the diagonal elements, i.e., $S = \{(i,j),\ i \neq j\}$. If the marginal totals are $x_{i\cdot} = x_{\cdot j} = 1$ for all $1 \leq i,j \leq 3$, there are two states in $\mathcal{F}(\{x_{i\cdot}\},\{x_{\cdot j}\},S)$ displayed as

$$\begin{array}{|ccc|} [0] & 1 & 0 \\ 0 & [0] & 1 \\ 1 & 0 & [0] \end{array} \quad \text{and} \quad \begin{array}{|ccc|} [0] & 0 & 1 \\ 1 & [0] & 0 \\ 0 & 1 & [0] \end{array}.$$

Here and hereafter we denote a structural zero cell by [0], in order to distinguish it from a sampling zero cell. To connect these two states, moves such as

$$\begin{array}{|ccc|} 0 & -1 & +1 \\ +1 & 0 & -1 \\ -1 & +1 & 0 \end{array}$$

are needed. In Smith *et al.* (1996), it is remarked without proofs that a connected chain can be constructed only by the moves such as (3) for $I \times I$ contingency tables with structural zeros as diagonal cells, when $I > 3$. In Section 7 we show that this statement is true when all the marginal totals are positive. But our concern in this paper is to list a common minimal set of moves that is needed to construct a connected chain for *arbitrary* values of the marginal totals, depending only on the size of the contingency tables and S, the configuration of the structural zero cells.

To describe the problem precisely, we give a definition of a *Markov basis* and its *minimality* as introduced in Diaconis and Sturmfels (1998). Let $\mathcal{F}_0(S)$ be the set of $I \times J$ contingency tables with non-structural zero cells as S and zero marginal totals

$$\mathcal{F}_0(S) = \left\{ \boldsymbol{Y} \;\middle|\; \sum_{j=1}^{J} y_{ij} = \sum_{i=1}^{I} y_{ij} = 0,\ y_{ij} \in \mathbb{Z},\ \text{and } y_{ij} = 0 \text{ for } (i,j) \notin S \right\},$$

where $\mathbb{Z}$ denotes the set of integers. The elements of $\mathcal{F}_0(S)$ are called *moves* on S.

Definition 1 *A Markov basis is a set $\mathcal{B} = \{\boldsymbol{B}_1,\ldots,\boldsymbol{B}_L\}$ of $I \times J$ contingency tables $\boldsymbol{B}_i \in \mathcal{F}_0(S)$, $i = 1\ldots,L$, such that, for any $\{x_{i\cdot}\},\{x_{\cdot j}\}$ and $\boldsymbol{X},\boldsymbol{X}' \in \mathcal{F}(\{x_{i\cdot}\},\{x_{\cdot j}\},S)$, there exist $A > 0$, $(\varepsilon_1,\boldsymbol{B}_{t_1}),\ldots,(\varepsilon_A,\boldsymbol{B}_{t_A})$ with $\varepsilon_s = \pm 1$, such that*

$$\boldsymbol{X}' = \boldsymbol{X} + \sum_{s=1}^{A} \varepsilon_s \boldsymbol{B}_{t_s} \quad \text{and} \quad \boldsymbol{X} + \sum_{s=1}^{a} \varepsilon_s \boldsymbol{B}_{t_s} \in \mathcal{F}(\{x_{i\cdot}\},\{x_{\cdot j}\},S) \text{ for } 1 \leq a \leq A.$$

A Markov basis $\mathcal{B}$ is minimal if no proper subset of $\mathcal{B}$ is a Markov basis.

4 Unique minimal basis for quasi-independence in two-way incomplete contingency tables

In this section, we derive a minimal Markov basis for $I \times J$ contingency tables with structural zeros. In the following, we assume that the level indices $i_1, i_2, \dots$ and $j_1, j_2, \dots$ are all distinct, i.e.,

$$i_m \neq i_n \text{ and } j_m \neq j_n \text{ for all } m \neq n. \tag{4}$$

The following definition gives a fundamental tool in this paper.

Definition 2 *A loop (or loop move) of degree r on S is an $I \times J$ contingency table*

$$M_r(i_1, \dots, i_r; j_1, \dots, j_r) \in \mathcal{F}_0(S), \quad 1 \leq i_1, \dots, i_r \leq I, \ 1 \leq j_1, \dots, j_r \leq J,$$

where $M_r(i_1, \dots, i_r; j_1, \dots, j_r)$ has the elements

$$\begin{aligned} m_{i_1 j_1} = m_{i_2 j_2} = \cdots = m_{i_{r-1} j_{r-1}} = m_{i_r j_r} = 1, \\ m_{i_1 j_2} = m_{i_2 j_3} = \cdots = m_{i_{r-1} j_r} = m_{i_r j_1} = -1, \end{aligned}$$

and all the other elements are zero. Specifically degree 2 loop $M_2(i_1, i_2; j_1, j_2)$ is called a basic move. The support of $M_r(i_1, \dots, i_r; j_1, \dots, j_r)$ is the set of its non-zero cells $\{(i_1, j_1), (i_1, j_2), \dots, (i_r, j_1)\}$.

Note that because of the notational convention (4), there is at most one $+1$ and -1 in each row and each column of a degree r loop. Loops constitute an essential subset of $\mathcal{F}_0(S)$. To see the role of the loops, we give the following lemma.

Lemma 1 *Let $\boldsymbol{Y} \in \mathcal{F}_0(S)$ be an $I \times J$ contingency table with some nonzero elements. Then $\boldsymbol{Y}$ can be expressed as a finite sum*

$$\boldsymbol{Y} = \sum_k a_k M_{r(k)}(i_{1(k)}, \dots, i_{r(k)}; j_{1(k)}, \dots, j_{r(k)}),$$

where a_k is a positive integer, $r(k) \leq \min\{I, J\}$ and there is no cancellation of signs in any cell.

The proof of this lemma is easy and omitted. Instead we show an example to clarify the meaning of this lemma. Let $\boldsymbol{Y} \in \mathcal{F}_0(S)$ be 4×5 contingency table expressed as follows.

3	−2	0	−2	1
−2	3	0	0	−1
−1	−1	2	0	0
0	0	−2	2	0

First notice that $y_{11} > 0$. Since $y_{1\cdot} = y_{\cdot 1} = 0$, there exists some $i, j > 1$ satisfying $y_{1j}, y_{i1} < 0$. In this case, we see that $y_{12}, y_{14}, y_{21}, y_{31} < 0$. Here, since $y_{22} > 0$, the four cells, $y_{11}, y_{12}, y_{22}, y_{21}$, are the nonzero elements of the basic move $M_2(1,2;1,2)$. Then we can subtract (twice) this basic move from $\boldsymbol{Y}$ as follows.

$$\boldsymbol{Y} = \begin{array}{|rrrrr|}\hline 2 & -2 & 0 & 0 & 0 \\ -2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline \end{array} + \begin{array}{|rrrrr|}\hline 1 & 0 & 0 & -2 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & -2 & 2 & 0 \\ \hline \end{array}$$

$$= \quad 2M_2(1,2;1,2) \quad + \quad \boldsymbol{Y}'$$

Note that there is no cancellation of signs in any cell and the remaining pattern, $\boldsymbol{Y}' = \boldsymbol{Y} - 2M_2(1,2;1,2)$, is again in $\mathcal{F}_0(S)$. Hence we can consider a further decomposition of $\boldsymbol{Y}'$. We observe that the six cells, $y'_{11}, y'_{14}, y'_{44}, y'_{43}, y'_{33}, y'_{31}$, are the locations of the nonzero elements of the degree 3 loop $M_3(1,4,3;1,4,3)$, and the remaining pattern is $M_4(1,4,3,2;5,4,3,2)$. Now the following decomposition of $\boldsymbol{Y}$ is obtained.

$$\boldsymbol{Y} = \begin{array}{|rrrrr|}\hline 2 & -2 & 0 & 0 & 0 \\ -2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline \end{array} + \begin{array}{|rrrrr|}\hline 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ \hline \end{array} + \begin{array}{|rrrrr|}\hline 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ \hline \end{array}$$

$$= \quad 2M_2(1,2;1,2) \quad + \quad M_3(1,4,3;1,4,3) \quad + \quad M_4(1,4,3,2;5,4,3,2)$$

It should be noted that this is not the only decomposition of $\boldsymbol{Y}$.

$$\boldsymbol{Y} = M_2(1,2;1,2) + M_2(1,2;5,2) + M_3(1,4,3;1,4,3) + M_4(1,4,3,2;1,4,3,2)$$

is another decomposition of $\boldsymbol{Y}$, satisfying the condition of Lemma 1. Lemma 1 describes the relation between Definition 2 and a Markov chain over $\mathcal{F}(\{x_{i\cdot}\},\{x_{\cdot j}\},S)$. Suppose $\boldsymbol{X},\boldsymbol{X}' \in \mathcal{F}(\{x_{i\cdot}\},\{x_{\cdot j}\},S)$. Then the difference $\boldsymbol{Y} = \boldsymbol{X} - \boldsymbol{X}'$ is in $\mathcal{F}_0(S)$. Hence to move from $\boldsymbol{X}$ to $\boldsymbol{X}'$, we can add a sequence of loops in Definition 2 to $\boldsymbol{X}$, without forcing negative entries on the way. In other words, a set of all the loops of degree $2,\dots,\min\{I,J\}$ constitute a trivial Markov basis.

From the definition, we have the relations

$$\begin{aligned} M_r(i_1,\dots,i_r;j_1,\dots,j_r) &= M_r(i_2,\dots,i_r,i_1;j_2,\dots,j_r,j_1) \\ &= -M_r(i_{r-1},i_{r-2},\dots,i_2,i_1,i_r;j_r,j_{r-1},\dots,j_2,j_1). \end{aligned}$$

Using these relations, we have $2r$ equivalent representations for a degree r loop. For example, a degree 4 loop used above

$$\begin{array}{|rrrrr|}\hline 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ \hline \end{array}$$

is expressed as either as

$$\begin{array}{rrr} M_4(1,4,3,2;5,4,3,2), & M_4(4,3,2,1;4,3,2,5), & M_4(3,2,1,4;3,2,5,4), \\ M_4(2,1,4,3;2,5,4,3), & -M_4(3,4,1,2;2,3,4,5), & -M_4(4,1,2,3;3,4,5,2), \\ -M_4(1,2,3,4;4,5,2,3), & -M_4(2,3,4,1;5,2,3,4). & \end{array}$$

Then if there is no structural zero cells, there are

$$\sum_{r=2}^{\min\{I,J\}} \binom{I}{r}\binom{J}{r}\frac{(r!)^2}{2r}$$

distinct loops in $I \times J$ contingency tables.

It is well known that the set of all basic moves constitutes a Markov basis for the case of no structural zero cells. Moreover, it is shown in Takemura and Aoki (2002) that this is the unique minimal Markov basis. Similarly in the presence of structural zeros, the set of loops is generally not a minimal basis. In this paper, we give an explicit characterization of the unique minimal Markov basis for an arbitrary configuration of structural zero cells. The next definition provides a key tool.

Definition 3 *A loop $M_r(i_1,\dots,i_r;j_1,\dots,j_r)$ is called df 1 if for all $a, 1 \le a \le r$, neither $R(i_1,\dots,i_{a-1},i_{a+1},\dots,i_r;j_1,\dots,j_r)$ nor $R(i_1,\dots,i_r;j_1,\dots,j_{a-1},j_{a+1},\dots,j_r)$ contains support of any loop on S of degree $2,\dots,r-1$, where*

$$R(i_1,\dots,i_r;j_1,\dots,j_r) = \{(i,j) \mid i \in \{i_1,\dots,i_r\}, j \in \{j_1,\dots,j_r\}\}.$$

Here the term *df* is intended as "degree of freedom". To clarify the meaning of this definition, we give an equivalent condition to Definition 3 in the following lemma.

Lemma 2 *$M_r(i_1,\dots,i_r;j_1,\dots,j_r)$ is df 1 if and only if $R(i_1,\dots,i_r;j_1,\dots,j_r)$ contains exactly two elements in S in every row and column.*

Proof. The case $r=2$ is obvious. Consider $r \ge 3$.

(Sufficiency) We argue by contradiction. By permuting the rows and columns, without loss of generality suppose that $M_r(1,\dots,r;1,\dots,r)$ is a degree r loop which does not satisfy the statement of the lemma. We also suppose $(1,a) \in S,\ 3 \le \exists a \le r$, without loss of generality. Then this loop is decomposed as

$$\begin{aligned} M_r(1,\dots,r;1,\dots,r) &= M_{r-a+2}(1,a,a+1,\dots,r;1,a,a+1,\dots,r) \\ &\quad + M_{a-1}(1,2,\dots,a-1;a,2,3,\dots,a-1). \end{aligned} \tag{5}$$

An example of $r=5, a=4$ is displayed as follows.

$$\begin{array}{|ccccc|}\hline +1 & -1 & [0] & 0 & [0] \\ [0] & +1 & -1 & [0] & [0] \\ [0] & [0] & +1 & -1 & [0] \\ [0] & [0] & [0] & +1 & -1 \\ -1 & [0] & [0] & [0] & +1 \\ \hline \end{array} = \begin{array}{|ccccc|}\hline +1 & 0 & [0] & -1 & [0] \\ [0] & 0 & 0 & [0] & [0] \\ [0] & [0] & 0 & 0 & [0] \\ [0] & [0] & [0] & +1 & -1 \\ -1 & [0] & [0] & [0] & +1 \\ \hline \end{array} + \begin{array}{|ccccc|}\hline 0 & -1 & [0] & +1 & [0] \\ [0] & +1 & -1 & [0] & [0] \\ [0] & [0] & +1 & -1 & [0] \\ [0] & [0] & [0] & 0 & 0 \\ 0 & [0] & [0] & [0] & 0 \\ \hline \end{array}$$

Here, the nonzero cells of the two loops, $M_{r-a+2}(1,a,a+1,\dots,r;1,a,a+1,\dots,r)$ and $M_{a-1}(1,2,\dots,a-1;a,2,3,\dots,a-1)$ overlap at $(1,a) \in S$ only. Therefore $R(1,a,a+1,\dots,r;1,a,a+1,\dots,r)$ and $R(1,\dots,a-1;2,\dots,a)$ contain the support $M_{r-a+2}(1,a,a+1,\dots,r;1,a,a+1,\dots,r)$ and $M_{a-1}(1,2,\dots,a-1;a,2,3,\dots,a-1)$, respectively, which are loops on S. Hence $M_r(1,\dots,r;1,\dots,r)$ is not df 1.

(Necessity) Suppose that $M_r(1,\dots,r;1,\dots,r)$ is a degree r loop such that $R(1,\dots,r;1,\dots,r)$ contains exactly two elements in S in every row and column. Without loss of generality, it is sufficient to show that $R(1,\dots,r-1;1,\dots,r)$ does not contain support of any loop on S. An example of $R(1,2,3,4;1,2,3,4,5)$ is displayed as follows.

m_{11}	m_{12}	[0]	[0]	[0]
[0]	m_{22}	m_{23}	[0]	[0]
[0]	[0]	m_{33}	m_{34}	[0]
[0]	[0]	[0]	m_{44}	m_{45}

From the assumption, $(1,1)$ is the only cell in S in $R(1,\dots,r-1;1)$, since $M_r(1,\dots,r;1,\dots,r)$ has exactly two nonzero elements there, i.e., $m_{11}=+1$ and $m_{r1}=-1$. Hence m_{11} is zero in any loop in $R(1,\dots,r-1;1,\dots,r)$. Moreover, by using the constraints $m_{1\cdot}=m_{\cdot 2}=m_{2\cdot}=\cdots=m_{r-1\cdot}=0$, it is shown that only the element of $\mathcal{F}_0(S)$ that can be contained in $R(1,\dots,r-1;1,\dots,r)$ is the zero contingency table. Q.E.D.

Lemma 2 describes the forms of the df 1 loops. The displays below are examples of df 1 loops of degree 2,3,4 in 4×5 contingency tables.

+1	−1	0	0	0
−1	+1	0	0	0
0	0	0	0	0
0	0	0	0	0

+1	−1	[0]	0	0
−1	[0]	+1	0	0
[0]	+1	−1	0	0
0	0	0	0	0

+1	−1	[0]	[0]	0
−1	[0]	+1	[0]	0
[0]	+1	[0]	−1	0
[0]	[0]	−1	+1	0

Obviously, every basic move is df 1. The term df 1 is motivated by the following consideration. Denote the positive and the negative part of a loop $M_r(i_1,\dots,i_r;j_1,\dots,j_r)$ as $M_r^+(i_1,\dots,i_r;j_1,\dots,j_r)$ and $M_r^-(i_1,\dots,i_r;j_1,\dots,j_r)$, respectively, i.e.,

$$m_{ij}^+=\max(m_{ij},0),\ \ m_{ij}^-=\max(-m_{ij},0).$$

Then

$$M_r(i_1,\dots,i_r;j_1,\dots,j_r)=M_r^+(i_1,\dots,i_r;j_1,\dots,j_r)-M_r^-(i_1,\dots,i_r;j_1,\dots,j_r). \tag{6}$$

Here, consider the set of contingency tables which have the same marginal totals and the configuration of S as $M_r^+(i_1,\dots,i_r;j_1,\dots,j_r)$ or $M_r^-(i_1,\dots,i_r;j_1,\dots,j_r)$, in other words,

$$\mathcal{F}(\{m_{i\cdot}^+\},\{m_{\cdot j}^+\},S)=\mathcal{F}(\{m_{i\cdot}^-\},\{m_{\cdot j}^-\},S).$$

Then we see that this set is a two-elements set with $M_r^+(i_1,\dots,i_r;j_1,\dots,j_r)$ and $M_r^-(i_1,\dots,i_r;j_1,\dots,j_r)$ being the only members.

Here we give our main theorem.

Theorem 1 *The set of df 1 loops of degree $2,\dots,\min\{I,J\}$ constitutes a unique minimal Markov basis for $I\times J$ contingency tables.*

Proof. We have already seen that the set of loops forms a Markov basis. We also note that every minimal Markov basis has to contain all loops of degrees $2, \dots, \min\{I, J\}$ on S. This is because, as we have seen in (6), df 1 loop is written as the difference of the two elements of $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$, which is exactly the two-elements set. Following the arguments of Takemura and Aoki (2002), we only need to verify that the set of the df 1 loops is itself a Markov basis. In order to show this, we argue by induction. We start from the trivial Markov basis consisting of all loops. We look at a non-df-1 loop of the highest degree. Below we show that this loop can be replaced by a combination of loops of smaller degrees, so that the resulting set is still a Markov basis. Then by induction on the highest degree of non-df-1 loops, it follows that the set of df 1 loops is a Markov basis.

In order to show that a non-df-1 loop of the highest degree can be replaced by combination of loops of lower degrees we again use the decomposition of loops that we have already seen. Suppose a Markov basis contains non-df-1 loops. Without loss of generality let $M_r(1, \dots, r; 1, \dots, r)$ be a non-df-1 loop of the highest degree and $(1, a) \in S, 3 \le \exists a \le r$. Then this loop is decomposed as (5). Here, the two loops, $M_{r-a+2}(1, a, a+1, \dots, r; 1, a, a+1, \dots, r)$ and $M_{a-1}(1, 2, \dots, a-1; a, 2, 3, \dots, a-1)$, overlap, i.e., have nonzero element in common position, only at $(1, a) \in S$. Since $(1, a)$ elements of these loops are -1 and $+1$, respectively, we can add or subtract these loops in an appropriate order to $\boldsymbol{X} \in \mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$ without forcing negative entries on the way, instead of adding or subtracting $M_r(1, \dots, r; 1, \dots, r)$ to $\boldsymbol{X}$. Therefore $M_r(1, \dots, r; 1, \dots, r)$ can be removed from the Markov basis and the remaining set is still a Markov basis. Q.E.D.

We clarify the last step of the above proof by an example. Let $\boldsymbol{X}$ and $\boldsymbol{X}'$ be

$$\boldsymbol{X} = \begin{array}{|ccccc|} \hline 0 & 1 & [0] & 0 & [0] \\ [0] & 0 & 1 & [0] & [0] \\ [0] & [0] & 0 & 1 & [0] \\ [0] & [0] & [0] & 0 & 1 \\ 1 & [0] & [0] & [0] & 0 \\ \hline \end{array}\,, \quad \boldsymbol{X}' = \begin{array}{|ccccc|} \hline 1 & 0 & [0] & 0 & [0] \\ [0] & 1 & 0 & [0] & [0] \\ [0] & [0] & 1 & 0 & [0] \\ [0] & [0] & [0] & 1 & 0 \\ 0 & [0] & [0] & [0] & 1 \\ \hline \end{array}\,.$$

$\boldsymbol{X}$ and $\boldsymbol{X}'$ are in the same $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$ and the difference $\boldsymbol{X}' - \boldsymbol{X}$ is expressed as a non-df-1 loop,

$$M_5(1,2,3,4,5;1,2,3,4,5) = \begin{array}{|ccccc|} \hline +1 & -1 & [0] & 0 & [0] \\ [0] & +1 & -1 & [0] & [0] \\ [0] & [0] & +1 & -1 & [0] \\ [0] & [0] & [0] & +1 & -1 \\ -1 & [0] & [0] & [0] & +1 \\ \hline \end{array}\,.$$

However we have already seen the decomposition

$$\begin{aligned} M_5(1,2,3,4,5;1,2,3,4,5) &= \begin{array}{|ccccc|} \hline +1 & 0 & [0] & -1 & [0] \\ [0] & 0 & 0 & [0] & [0] \\ [0] & [0] & 0 & 0 & [0] \\ [0] & [0] & [0] & +1 & -1 \\ -1 & [0] & [0] & [0] & +1 \\ \hline \end{array} + \begin{array}{|ccccc|} \hline 0 & -1 & [0] & +1 & [0] \\ [0] & +1 & -1 & [0] & [0] \\ [0] & [0] & +1 & -1 & [0] \\ [0] & [0] & [0] & 0 & 0 \\ 0 & [0] & [0] & [0] & 0 \\ \hline \end{array} \\ &= \qquad M_3(1,4,5;1,4,5) \qquad + \qquad M_3(1,2,3;4,2,3). \end{aligned}$$

Seeing that the $(1,4)$ element of $M_3(1,2,3;4,2,3)$ is positive, it follows $\boldsymbol{X} + M_3(1,2,3;4,2,3) \in \mathcal{F}(\{x_{i\cdot}\},\{x_{\cdot j}\},S)$ as

$$\boldsymbol{X} + M_3(1,2,3;4,2,3) = \begin{array}{|ccccc|} \hline 0 & 0 & [0] & 1 & [0] \\ [0] & 1 & 0 & [0] & [0] \\ [0] & [0] & 1 & 0 & [0] \\ [0] & [0] & [0] & 0 & 1 \\ 1 & [0] & [0] & [0] & 0 \\ \hline \end{array}.$$

Now, seeing that $(1,4),(4,5),(5,1)$ elements are positive, we can add $M_3(1,4,5;1,4,5)$ to $\boldsymbol{X} + M_3(1,2,3;4,2,3)$ and obtain $\boldsymbol{X}'$. Hence it is shown that the non-df-1 loop $M_5(1,2,3,4,5;1,2,3,4,5)$ is redundant. It should also be noted that the two loops, $M_3(1,4,5;1,4,5)$ and $M_3(1,2,3;4,2,3)$, are both df 1 and hence cannot be removed.

Example. Comparison of the minimal basis and the reduced Gröbner basis. Consider 6×6 contingency tables of the following form.

$$\begin{array}{|cccccc|} \hline [0] & x_{12} & x_{13} & [0] & [0] & x_{16} \\ x_{21} & [0] & x_{23} & x_{24} & [0] & [0] \\ x_{31} & x_{32} & [0] & [0] & x_{35} & [0] \\ [0] & [0] & x_{43} & [0] & x_{45} & x_{46} \\ x_{51} & [0] & [0] & x_{54} & [0] & x_{56} \\ [0] & x_{62} & [0] & x_{64} & x_{65} & [0] \\ \hline \end{array}$$

By the algebraic algorithm described in Diaconis and Sturmfels (1998), we calculated the reduced Gröbner basis using the degree reverse lexicographical ordering. The result was composed of 3 basic moves, 20 degree 3 loops, 10 degree 4 loops and 3 degree 5 loops. The following is a list of these loops.

$M_2(1,4;6,3)$	$M_2(2,5;4,1)$	$M_2(3,6;5,2)$
$M_3(1,2,3;2,3,1)$	$M_3(1,2,5;6,3,1)$	$M_3(1,2,5;6,3,4)$
$M_3(1,3,4;6,2,5)$	$M_3(1,3,5;6,2,1)$	$M_3(1,4,3;2,3,5)$
$M_3(1,6,2;3,2,4)$	$M_3(1,6,4;3,2,5)$	$M_3(1,6,4;6,2,5)$
$M_3(1,6,5;6,2,4)$	$M_3(2,3,6;4,1,2)$	$M_3(2,4,3;1,3,5)$
$M_3(2,5,4;3,1,6)$	$M_3(2,5,4;3,4,6)$	$M_3(2,6,3;1,4,5)$
$M_3(2,6,4;3,4,5)$	$M_3(3,5,4;5,1,6)$	$M_3(3,5,6;5,1,4)$
$M_3(3,6,5;1,2,4)$	$M_3(4,5,6;5,6,4)$	
$M_4(1,2,3,4;6,3,1,5)$	$M_4(1,2,6,3;2,3,4,5)$	$M_4(1,3,2,5;6,2,1,4)$
$M_4(1,3,5,4;3,2,1,6)$	$M_4(1,5,6,3;2,6,4,5)$	$M_4(1,6,5,4;3,2,4,6)$
$M_4(1,6,5,2;3,2,4,1)$	$M_4(2,3,6,4;3,1,2,5)$	$M_4(2,5,4,3;1,4,6,5)$
$M_4(3,6,4,5;1,2,5,6)$		
$M_5(1,3,2,5,4;3,2,1,4,6)$	$M_5(1,3,2,6,4;3,2,1,4,5)$	$M_5(1,3,5,6,4;3,2,1,4,5)$

This list is very confusing and we cannot recognize the structure of the basis at first sight. One reason for the difficulty is that the above list is not minimal. (Note that the reduced Gröbner basis may not be a minimal basis.) It can be easily checked that the loops of degree 4 and 5 are not df 1. On the other hand, all the 20 loops of degree 3 are df 1. Hence from Theorem 1, the above 3 basic moves and 20 degree 3 loops constitute the unique minimal Markov basis.

In Section 5, a simple algorithm to list all the elements of the unique minimal Markov basis is given.

5 Algorithms for enumerating elements of a minimal basis

In this section, we discuss how to list all the elements of the unique minimal basis. As we have seen, the elements of the unique minimal Markov basis have a simple structure described in Lemma 2. Considering this structure, we have an explicit form of a minimal basis for some typical situations, which play important roles in applications. We consider these special cases first and then consider general cases.

Separable tables Separability is one of the most important concepts for analyzing incomplete contingency tables. The definition of the separability is as follows (Mantel, 1970, Bishop *et al.*, 1975). In a two-way contingency table two cells are *associated* if they do not contain structural zeros and if they are either in the same row or the same column. A set of non-structural zero cells is *connected* if every pairs of cells can be linked by a chain of cells, any two consecutive members of which must be associated. Finally, an incomplete two-way table is connected if its non-structural zero cells form a connected set. An incomplete table that is not connected is said to be *separable*. Separable two-way contingency tables can be rearranged to a block diagonal form with connected subtables by permuting the rows and columns. Table 1 is an example of separable table from Harris (1910). By permuting the rows and columns, we see that this table is separable with exactly two connected subtables as displayed in Table 2.

Table 1: An example of a separable table:
Relationship between radial asymmetry and locular composition in *Staphylea* (Series A of Harris, 1910)

locular	coefficient of radial asymmetry								
composition	0.00	0.47	0.82	0.94	1.25	1.41	1.63	1.70	1.89
3 even, 0 odd	462	[0]	[0]	130	[0]	[0]	2	[0]	1
2 even, 1 odd	[0]	614	138	[0]	21	14	[0]	1	[0]
1 even, 2 odd	[0]	443	95	[0]	22	8	[0]	5	[0]
0 even, 3 odd	103	[0]	[0]	35	[0]	[0]	1	[0]	0

We see easily that the minimal Markov basis for this example consists of basic moves only. This is obvious from the fact that the two connected subtables do not contain structural zero cells respectively. When the connected subtables contain some structural zero cells, the minimal Markov basis for the whole table is a union of the minimal Markov bases for these subtables.

Table 2: Data from Table 1 after rearrangement of rows and columns

locular composition	coefficient of radial asymmetry 0.00	0.94	1.63	1.89	0.47	0.82	1.25	1.41	1.70
3 even, 0 odd	462	130	2	1	[0]	[0]	[0]	[0]	[0]
0 even, 3 odd	103	35	1	0	[0]	[0]	[0]	[0]	[0]
2 even, 1 odd	[0]	[0]	[0]	[0]	614	138	21	14	1
1 even, 2 odd	[0]	[0]	[0]	[0]	443	95	22	8	5

For example, the minimal Markov basis for the following separable 6×7 contingency table

$$\begin{array}{|ccccccc|}\hline x_{11} & x_{12} & x_{13} & [0] & [0] & [0] & [0] \\ [0] & [0] & x_{23} & x_{24} & [0] & [0] & [0] \\ x_{31} & x_{32} & [0] & x_{34} & [0] & [0] & [0] \\ [0] & [0] & [0] & [0] & x_{45} & x_{46} & [0] \\ [0] & [0] & [0] & [0] & [0] & x_{56} & x_{57} \\ [0] & [0] & [0] & [0] & x_{65} & [0] & x_{67} \\ \hline \end{array}$$

is the union of the minimal Markov basis for two subtables,

$$\begin{array}{|cccc|}\hline x_{11} & x_{12} & x_{13} & [0] \\ [0] & [0] & x_{23} & x_{24} \\ x_{31} & x_{32} & [0] & x_{34} \\ \hline \end{array} \quad \text{and} \quad \begin{array}{|ccc|}\hline x_{45} & x_{46} & [0] \\ [0] & x_{56} & x_{57} \\ x_{65} & [0] & x_{67} \\ \hline \end{array}.$$

We see it is $\{M_2(1,3;1,2), M_3(1,2,3;1,3,4), M_3(1,2,3;2,3,4), M_3(4,5,6;5,6,7)\}$.

Block triangular tables Another typical situation is that an incomplete table is in *block triangular* form, i.e., after suitable permutation of rows and columns, $(i,j) \notin S$ implies $(k,l) \notin S$ for all $k \geq i$ and $l \geq j$ (Goodman, 1968, Bishop *et al.*, 1975). The following tables are examples of block triangular tables.

$$\begin{array}{|cccc|}\hline x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & [0] \\ x_{31} & x_{32} & [0] & [0] \\ x_{41} & [0] & [0] & [0] \\ \hline \end{array} \quad \begin{array}{|cccc|}\hline [0] & [0] & x_{13} & x_{14} \\ [0] & [0] & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \\ \hline \end{array} \quad \begin{array}{|cccc|}\hline x_{11} & x_{12} & x_{13} & x_{14} \\ [0] & x_{22} & x_{23} & x_{24} \\ [0] & x_{32} & x_{33} & x_{34} \\ [0] & [0] & x_{43} & x_{44} \\ [0] & [0] & x_{53} & x_{54} \\ \hline \end{array}$$

Table 3 shows an example of a block triangle contingency table from Bishop and Fienberg (1969). The minimal Markov bases for these tables are simple, i.e., the set of basic moves constitutes the minimal Markov basis. Hence a Metropolis-Hasting sampling can be constructed simply by choosing pairs of rows and columns, which intersect at non-structural zero cells.

Square tables with diagonal elements being structural zeros There are many situations that the contingency tables are square and all the diagonal elements are structural zero cells. Table 4 is an example of such tables. It is obvious that the minimal Markov basis for

Table 3: An example of a block triangular table:
Initial and final ratings on disability of stroke patients

initial state	final state A	B	C	D	E
E	11	23	12	15	8
D	9	10	4	1	[0]
C	6	4	4	[0]	[0]
B	4	5	[0]	[0]	[0]
A	5	[0]	[0]	[0]	[0]

Source: Bishop and Fienberg (1969).

Table 4: An example of a square table
with diagonal elements being structural zeros

active participant	passive participant R	S	T	U	V	W
R	[0]	1	5	8	9	0
S	29	[0]	14	46	4	0
T	0	0	[0]	0	0	0
U	2	3	1	[0]	28	2
V	0	0	0	0	[0]	1
W	9	25	4	6	13	[0]

Source: Ploog (1967).

such tables contains degree 3 loops which correspond to every triplet of the structural zeros. For examples, degree 3 loops such as

$$\begin{array}{|cccccc|} \hline [0] & -1 & +1 & 0 & 0 & 0 \\ +1 & [0] & -1 & 0 & 0 & 0 \\ -1 & +1 & [0] & 0 & 0 & 0 \\ 0 & 0 & 0 & [0] & 0 & 0 \\ 0 & 0 & 0 & 0 & [0] & 0 \\ 0 & 0 & 0 & 0 & 0 & [0] \\ \hline \end{array} \quad \text{or} \quad \begin{array}{|cccccc|} \hline [0] & 0 & 0 & 0 & 0 & 0 \\ 0 & [0] & 0 & 0 & +1 & -1 \\ 0 & 0 & [0] & 0 & 0 & 0 \\ 0 & 0 & 0 & [0] & 0 & 0 \\ 0 & -1 & 0 & 0 & [0] & +1 \\ 0 & +1 & 0 & 0 & -1 & [0] \\ \hline \end{array}$$

are needed to construct a connected Markov chain. It is seen that for $I \times I$ contingency tables, there are $\binom{I}{2}\binom{I-2}{2}$ basic moves and $\binom{I}{3}$ df 1 degree 3 loops in the minimal Markov basis.

For such types of contingency tables, the hypothesis of quasi-symmetry is also of interest in many situations (Smith *et al.*, 1996). We also derive the unique minimal Markov basis for the case of a quasi-symmetry hypothesis in Appendix.

General incomplete tables We have seen some typical situations which frequently appear in applications. Now we give some rules and algorithms to list all the elements in the minimal

Markov basis for arbitrary configuration of structural zeros. Table 5 is an example of incomplete tables which cannot be categorized as any type discussed above. In view of Lemma 2, it may

Table 5: Classification of Purum marriages

	Sib of husband				
Sib of wife	Marrim	Makan	Parpa	Thao	Kheyang
Marrim	[0]	5	17	[0]	6
Makan	5	[0]	0	16	2
Parpa	[0]	2	[0]	10	11
Thao	10	[0]	[0]	[0]	9
Kheyang	6	20	8	0	1

Source: White (1963), based on data of Das (1945).

be easy to list all the elements of the unique minimal basis for many situations, especially when there are only a few structural zero cells.

It is also easy to obtain an upper bound of the degree of loops which is in the minimal basis, simply by counting the number of structural zero cells in each row and column. It should be noted that if the minimal Markov basis contains a degree r loop, then there are at least r rows and r columns which have $r-2$ structural zero cells. Hence for Table 5, we see that the degree of the loops contained in the minimal Markov basis is at most 3. For these types of contingency tables, simple enumeration algorithms may be effective. Indeed, for Table 5, we see that $M_3(1,2,3;2,3,4)$ is the only degree 3 loop in the minimal basis by considering every triplet of structural zero cells located in distinct rows and columns.

In general, we can make use of the following recursive algorithm to list all the elements in the minimal basis. This algorithm works especially well in some situations that the contingency table is *near* separable, or *semi-separable* (Mantel, 1970), i.e., when the table can be made separable into two or more connected subtables by the removal of a single row or a single column. The following tables are examples of semi-separable tables.

(i)

x_{11}	x_{12}	[0]	[0]	[0]
x_{21}	x_{22}	[0]	[0]	x_{25}
[0]	[0]	x_{33}	x_{34}	x_{35}
[0]	[0]	x_{43}	x_{44}	[0]
[0]	[0]	x_{53}	x_{54}	x_{55}

(ii)

x_{11}	x_{12}	[0]	[0]	[0]
x_{21}	x_{22}	[0]	[0]	[0]
[0]	[0]	x_{33}	x_{34}	x_{35}
[0]	[0]	x_{43}	x_{44}	[0]
x_{51}	x_{52}	[0]	x_{54}	x_{55}

We see that these tables are made to be separable by removal of *(i)* the row 2 or the column 5 and *(ii)* the row 5, respectively. Now we give a simple recursive algorithm.

```
Input:  I_0 = {1,...,I}, J_0 = {1,...,J}, S
Output:  elements of a minimal basis

ListMoves(I_0; J_0)
{
```

```
    Choose i* ∈ I_0 and J* = {j | (i*, j) ∈ S};
    List df 1 moves which have ±1 elements in R(i*; J*);
    ListMoves(I_0 − {i*}; J_0);
}
```

To illustrate the meanings of this algorithm, we reanalyze the 6×6 example discussed in Section 4, which is displayed below.

	1	2	3	4	5	6
1	[0]	x_{12}	x_{13}	[0]	[0]	x_{16}
2	x_{21}	[0]	x_{23}	x_{24}	[0]	[0]
3	x_{31}	x_{32}	[0]	[0]	x_{35}	[0]
4	[0]	[0]	x_{43}	[0]	x_{45}	x_{46}
5	x_{51}	[0]	[0]	x_{54}	[0]	x_{56}
6	[0]	x_{62}	[0]	x_{64}	x_{65}	[0]

It has been already mentioned that the minimal Markov basis for this table is a set of basic moves and 20 degree 3 loops. To see this, first we choose $i^* = 1$ and hence $J^* = \{2, 3, 6\}$. We also denote $\widetilde{I} = I_0 - \{i^*\} = \{2, 3, 4, 5, 6\}$, $\widetilde{J} = J_0 - J^* = \{1, 4, 5\}$.

		$\widetilde{J}$			J^*		
		1	4	5	2	3	6
i^*	1	[0]	[0]	[0]	x_{12}	x_{13}	x_{16}
	2	x_{21}	x_{24}	[0]	[0]	x_{23}	[0]
	3	x_{31}	[0]	x_{35}	x_{32}	[0]	[0]
$\widetilde{I}$	4	[0]	[0]	x_{45}	[0]	x_{43}	x_{46}
	5	x_{51}	x_{54}	[0]	[0]	[0]	x_{56}
	6	[0]	x_{64}	x_{65}	x_{62}	[0]	[0]

Next step of the algorithm is to list all df 1 loops which have ± 1 elements in $R(i^*; J^*)$. It is easy to do this because such loop has exactly one $+1$ and one -1 both in $R(i^*; J^*)$ and $R(\widetilde{I}; J^*)$. Hence we can list such loops by listing all pairs of columns in J^*. In this case,

- if select $(2, 3)$ from J^* then $M_3(1, 2, 3; 2, 3, 1), M_3(1, 6, 2; 3, 2, 4), M_3(1, 4, 3; 2, 3, 5)$ and $M_3(1, 6, 4; 3, 2, 5)$ are listed,
- if select $(2, 6)$ from J^* then $M_3(1, 3, 4; 6, 2, 5), M_3(1, 3, 5; 6, 2, 1), M_3(1, 6, 4; 6, 2, 5)$ and $M_3(1, 6, 5; 6, 2, 4)$ are listed,
- if select $(3, 6)$ from J^* then $M_2(1, 4; 6, 3), M_3(1, 2, 5; 6, 3, 1)$ and $M_3(1, 2, 5; 6, 3, 4)$ are listed.

These are all the df 1 loops which have ± 1 in this $i^* = 1$-th row. Now all we have to consider is the subtable $R(\widetilde{I}; J_0)$, to which we can apply the similar procedure, and finally the solution described in Section 4 is given.

In general case, it is effective to select i^* so that there are as many structural zero cells in $R(i^*; J_0)$ as possible. Hence i^* should be chosen as

$$i^* = \arg\max_i \#\{j \mid (i^*, j) \notin S\}.$$

However, if the table is semi-separable, it is also effective to select i^* so that the remaining table $R(\widetilde{I}; J_0)$ becomes separable.

6 Computational example

Using the Markov basis obtained above, we can perform various tests by the Monte Carlo method. In this section, we show an example of testing the hypothesis of quasi-independence for a given data set. Table 6 shows a data collected by Vidmar (1972) for discovering the possible effects on decision making of limiting the number of alternatives available to the members of a jury panel. This is a 4×7 contingency table which has 9 structural zero cells. The degrees

Table 6: Effects of decision alternatives on the verdicts and social perceptions of simulated jurors

	condition						
alternative	1	2	3	4	5	6	7
first-degree	11	[0]	[0]	2	7	[0]	2
second-degree	[0]	20	[0]	22	[0]	11	15
manslaughter	[0]	[0]	22	[0]	16	13	5
not guilty	13	4	2	0	1	0	2

Source: Vidmar (1972).

of freedom for testing quasi-independence is 9. The maximum likelihood estimate under the hypothesis of quasi-independence is calculated by iterative method as displayed in Table 7. See Bishop *et al.*(1975) for maximum likelihood estimation of incomplete tables.

Table 7: Maximum likelihood estimate for Table 6

	condition						
alternative	1	2	3	4	5	6	7
first-degree	14.05	[0]	[0]	2.61	3.64	[0]	1.70
second-degree	[0]	21.93	[0]	19.55	[0]	13.75	12.77
manslaughter	[0]	[0]	20.95	[0]	17.78	8.95	8.32
not guilty	9.95	2.07	3.05	1.84	2.58	1.30	1.21

As the discrepancy measure from the hypothesis of quasi-independence, we use the likelihood-ratio statistic

$$G^2 = 2 \sum_S x_{ij} \log \frac{x_{ij}}{\hat{m}_{ij}},$$

where $\hat{m}_{ij}$ is the MLE of the expectation parameter m_{ij}. The observed value of G^2 is 18.816 and the corresponding asymptotic p value is 0.0268 from the asymptotic distribution χ^2_9.

To perform the Markov chain Monte Carlo method, first we obtain the minimal Markov basis. From the considerations in the above sections, we see easily that a set of basic moves and a degree 3 loop $M_3(1, 2, 3; 5, 4, 6)$ constitute the unique minimal Markov basis. Using this basis, we construct a connected chain, which is modified so as to have the null distribution (2) as the stationary distribution by the Metropolis-Hasting procedure. The estimated exact p value is 0.0444, with estimated standard deviation 0.00052. (We use a batching method to obtain an estimate of variance. See Hastings, 1970, or Ripley, 1987.) Figure 1 shows a histogram of the Monte Carlo sampling generated from the exact distribution of the likelihood ratio statistic under the quasi-independence hypothesis, along with the corresponding asymptotic distribution χ^2_9. We see that the asymptotic distribution understates the probability that the test statistic

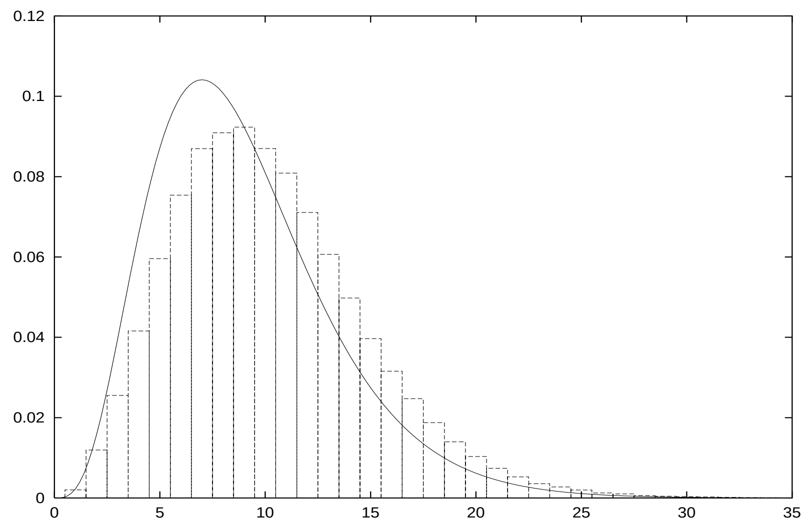

Figure 1: Asymptotic and Monte Carlo estimated exact distribution for the likelihood ratio statistic under the quasi-independence model

is greater than the observed value, and overemphasize the significance.

7 Basis reduction for the case of positive marginals

The minimality of the basis considered in the previous sections is based on the condition that the values of the marginal totals are *arbitrary*. However, for performing exact conditional tests to a given data set, we can assume without loss of generality that $x_{i\cdot}, x_{\cdot j} > 0$ for all i, j because all cell values in rows or columns with zero marginals are necessarily zeros and such rows or columns can be ignored in the conditional analysis.

Under the assumption of positive marginal totals, there may be cases, where some elements of the minimal basis obtained in Section 4 are not needed to construct a connected chain. Therefore it is worth investigating further reduction of the minimal basis to $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$ for fixed positive values of $x_{i\cdot}$ and $x_{\cdot j}$. From practical viewpoint it is important to give a sufficient condition on $x_{i\cdot}$ and $x_{\cdot j}$, such that a df 1 loop of degree $r \geq 3$ can be replaced by a series of basic moves. A simple example is the following 3×4 contingency table.

[0]	x_{12}	x_{13}	x_{14}
x_{21}	[0]	x_{23}	x_{24}
x_{31}	x_{32}	[0]	x_{34}

From the considerations of Section 4, we know that the 4 loops, $M_2(1,3;2,4)$, $M_2(1,2;3,4)$, $M_2(2,3;1,4)$ and $M_3(1,3,2;3,2,1)$, constitute the unique minimal Markov basis. They are displayed as follows.

$$\begin{array}{|cccc|}\hline 0 & +1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & +1 \\ \hline\end{array}, \begin{array}{|cccc|}\hline 0 & 0 & +1 & -1 \\ 0 & 0 & -1 & +1 \\ 0 & 0 & 0 & 0 \\ \hline\end{array}, \begin{array}{|cccc|}\hline 0 & 0 & 0 & 0 \\ +1 & 0 & 0 & -1 \\ -1 & 0 & 0 & +1 \\ \hline\end{array}, \begin{array}{|cccc|}\hline 0 & -1 & +1 & 0 \\ +1 & 0 & -1 & 0 \\ -1 & +1 & 0 & 0 \\ \hline\end{array}.$$

However, under the assumption that all the marginal totals are positive, it is shown that $M_3(1,3,2;3,2,1)$ is not needed to construct a connected chain. To see this, suppose 3×4 tables $\boldsymbol{X}, \boldsymbol{X}'$ are in $\mathcal{F}(\{x_{i\cdot}\},\{x_{\cdot j}\}, S=\{(1,1),(2,2),(3,3)\})$ and $\boldsymbol{X}'-\boldsymbol{X}=M_3(1,3,2;3,2,1)$. In this case, two states $\boldsymbol{X}$ and $\boldsymbol{X}'$ are mutually reachable by adding or subtracting $M_3(1,3,2;3,2,1)$. These two states can be written as

$$\boldsymbol{X}=\begin{array}{|cccc|}\hline [0] & a_1+1 & a_2 & a_3 \\ a_4 & [0] & a_5+1 & a_6 \\ a_7+1 & a_8 & [0] & a_9 \\ \hline\end{array} \text{ and } \boldsymbol{X}'=\begin{array}{|cccc|}\hline [0] & a_1 & a_2+1 & a_3 \\ a_4+1 & [0] & a_5 & a_6 \\ a_7 & a_8+1 & [0] & a_9 \\ \hline\end{array}$$

where $a_1,\ldots,a_9\in\mathbb{N}$. Here an important point is that at least one of a_3, a_6, a_9 is positive because $x_{\cdot 4}>0$. Then we can add or subtract the above 3 basic moves one by one in appropriate order to $\boldsymbol{X}$ to reach $\boldsymbol{X}'$. We show this procedure in the following when $a_3\geq 1$.

$$\boldsymbol{X}=\begin{array}{|cccc|}\hline [0] & a_1+1 & a_2 & a_3(\geq 1) \\ a_4 & [0] & a_5+1 & a_6 \\ a_7+1 & a_8 & [0] & a_9 \\ \hline\end{array}$$

$$\underset{+M_2(1,2;3,4)}{\longrightarrow}\begin{array}{|cccc|}\hline [0] & a_1+1 & a_2+1 & a_3-1 \\ a_4 & [0] & a_5 & a_6+1 \\ a_7+1 & a_8 & [0] & a_9 \\ \hline\end{array}$$

$$\underset{+M_2(2,3;1,4)}{\longrightarrow}\begin{array}{|cccc|}\hline [0] & a_1+1 & a_2+1 & a_3-1 \\ a_4+1 & [0] & a_5 & a_6 \\ a_7 & a_8 & [0] & a_9+1 \\ \hline\end{array}$$

$$\underset{-M_2(1,3;2,4)}{\longrightarrow}\begin{array}{|cccc|}\hline [0] & a_1 & a_2+1 & a_3 \\ a_4+1 & [0] & a_5 & a_6 \\ a_7 & a_8+1 & [0] & a_9 \\ \hline\end{array}=\boldsymbol{X}'$$

The above consideration gives a decomposition of a df 1 degree r loop into r basic moves, by using one additional row or column. If this row or column is known to contain at least one positive cell, a connected chain can simply be constructed by using basic moves, instead of using this degree r loop.

We summarize the above consideration in the following lemma.

Lemma 3 *Under the assumption of positive marginals, a df 1 loop* $M_r(i_1,\ldots,i_r;j_1,\ldots,j_r)$, $r\geq 3$, *can be replaced by a series of basic moves, if one of the following conditions is satisfied.*

(a) There exists $i^*\neq i_k, 1\leq k\leq r$, *such that* $(i^*,j_k)\in S$ *for all* $k=1,\ldots,r$ *and* $x_{i^*\cdot}>\sum_{j\in A} x_{\cdot j}$

where $A = \{j \mid j \neq j_k, 1 \leq k \leq r$ *and* $x_{i^*j} \in S\}$.

(b) There exists $j^* \neq j_k, 1 \leq k \leq r$, *such that* $(i_k, j^*) \in S$ *for all* $k = 1, \ldots, r$ *and* $x_{\cdot j^*} > \sum_{i \in A} x_{i \cdot}$

where $A = \{i \mid i \neq i_k, 1 \leq k \leq r$ *and* $x_{ij^*} \in S\}$.

Proof. It is sufficient to show the lemma for the case (a). It holds that

$$M_r(i_1, \ldots, i_r; j_1, \ldots, j_r) = \sum_{t=1}^{r-1} M_2(i_t, i^*; i_t, i_{t+1}) + M_2(i_r, i^*; i_r, 1). \tag{7}$$

It should be noted that the r basic moves of the right hand side constitute a cycle, of which the consecutive two basic moves have only one nonzero cell in common. Moreover, at least one of $x_{i^*j_1}, x_{i^*j_2}, \ldots, x_{i^*j_r}$ have to be positive for all the elements in $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$ from the positiveness of $x_{i^*\cdot}, x_{\cdot j}$. Hence we can add or subtract these r basic moves one by one to or from $\boldsymbol{X} \in \mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$ without forcing negative entries on the way, instead of using $M_r(i_1, \ldots, i_r; j_1, \ldots, j_r)$. Q.E.D.

To illustrate the argument of this lemma, let $M_4(1,2,3,4;1,2,3,4)$ be df 1 loop and and for example let $i^* = 5$ satisfy the condition (a). Then the decomposition (7) is written as

$$M_4(1,2,3,4;1,2,3,4) = M_2(1,5;1,2) + M_2(2,5;2,3) + M_2(3,5;3,4) + M_2(4,5;4,1). \tag{8}$$

In the case of $I = 5, J = 4$, this is displayed as

$$\begin{array}{|cccc|}\hline +1 & -1 & [0] & [0] \\ [0] & +1 & -1 & [0] \\ [0] & [0] & +1 & -1 \\ -1 & [0] & [0] & +1 \\ 0 & 0 & 0 & 0 \\ \hline\end{array} = \begin{array}{|cccc|}\hline +1 & -1 & [0] & [0] \\ [0] & 0 & 0 & [0] \\ [0] & [0] & 0 & 0 \\ 0 & [0] & [0] & 0 \\ -1 & +1 & 0 & 0 \\ \hline\end{array} + \begin{array}{|cccc|}\hline 0 & 0 & [0] & [0] \\ [0] & +1 & -1 & [0] \\ [0] & [0] & 0 & 0 \\ 0 & [0] & [0] & 0 \\ 0 & -1 & +1 & 0 \\ \hline\end{array}$$

$$+ \begin{array}{|cccc|}\hline 0 & 0 & [0] & [0] \\ [0] & 0 & 0 & [0] \\ [0] & [0] & +1 & -1 \\ 0 & [0] & [0] & 0 \\ 0 & 0 & -1 & +1 \\ \hline\end{array} + \begin{array}{|cccc|}\hline 0 & 0 & [0] & [0] \\ [0] & 0 & 0 & [0] \\ [0] & [0] & 0 & 0 \\ -1 & [0] & [0] & +1 \\ +1 & 0 & 0 & -1 \\ \hline\end{array}.$$

It should be noted that, in the above example, $A = \emptyset$ in the condition (a) and hence at least one of $x_{51}, \ldots, x_{54}$ must be positive from $x_{5\cdot} > 0$. Now consider adding $M_4(1,2,3,4;1,2,3,4)$ to some $\boldsymbol{X} \in \mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$. If $\boldsymbol{X} + M_4(1,2,3,4;1,2,3,4)$ is again in $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$, at least 4 entries of $\boldsymbol{X}$, $x_{12}, x_{23}, x_{34}, x_{41}$, must be positive. Then if at least one of $x_{51}, \ldots, x_{54}$ is positive, we can add some basic moves in the right hand side of (8). Suppose $x_{52} > 0$, for example, then $\boldsymbol{X} + M_2(2,5;2,3) \in \mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$ holds. Now the $(5,3)$ element of $\boldsymbol{X} + M_2(2,5;2,3)$ is positive, hence next basic move, $M_2(3,5;3,4)$, can be added without forcing negative entries. In the similar way, we can add all the basic moves in the right hand side of (7) without forcing negative entries and df 1 loop $M_r(i_1, \ldots, i_r; j_1, \ldots, j_r)$ is shown to be redundant.

We now consider another case not covered by Lemma 3. The following display is df 1 degree 3 loop $\boldsymbol{M}_3(1,2,3,2,3,1)$ in 5×5 square contingency table with diagonal elements being structural zeros.

$$
\begin{array}{|ccccc|}
\hline
[0] & +1 & -1 & 0 & 0 \\
-1 & [0] & +1 & 0 & 0 \\
+1 & -1 & [0] & 0 & 0 \\
0 & 0 & 0 & [0] & 0 \\
0 & 0 & 0 & 0 & [0] \\
\hline
\end{array}
$$

Of course, we can decompose this loop to 3 basic moves as directed in Lemma 3 as

$$
\boldsymbol{M}_3(1,2,3;2,3,1) = \begin{array}{|ccccc|}
\hline
[0] & +1 & -1 & 0 & 0 \\
0 & [0] & 0 & 0 & 0 \\
0 & 0 & [0] & 0 & 0 \\
0 & -1 & +1 & [0] & 0 \\
0 & 0 & 0 & 0 & [0] \\
\hline
\end{array} + \begin{array}{|ccccc|}
\hline
[0] & 0 & 0 & 0 & 0 \\
-1 & [0] & +1 & 0 & 0 \\
0 & 0 & [0] & 0 & 0 \\
+1 & 0 & -1 & [0] & 0 \\
0 & 0 & 0 & 0 & [0] \\
\hline
\end{array} + \begin{array}{|ccccc|}
\hline
[0] & 0 & 0 & 0 & 0 \\
0 & [0] & 0 & 0 & 0 \\
+1 & -1 & [0] & 0 & 0 \\
-1 & +1 & 0 & [0] & 0 \\
0 & 0 & 0 & 0 & [0] \\
\hline
\end{array}.
$$

However, it is not guaranteed that we can apply some of the above basic moves to every state which satisfies either of $x_{12}, x_{23}, x_{31} > 0$ or $x_{13}, x_{21}, x_{32} > 0$, because we cannot exclude the possibility that $x_{41} = x_{42} = x_{43} = 0$ even under the assumption of positive marginals. To deal with such cases, we give another decomposition of df 1 degree r loop into $r+1$ basic moves in the next lemma.

Lemma 4 *Under the assumption of positive marginals, a df 1 loop* $M_r(i_1,\ldots,i_r;j_1,\ldots,j_r)$, $r \geq 3$, *can be replaced by a series of basic moves, if one of the following conditions is satisfied.*

(a) There exists $i^* \neq i_k, 1 \leq k \leq r$, *such that* $(i^*, j_k) \in S$ *for all* $k = 1,\ldots,r$ *and for all* j *such that* $(i^*, j) \in S, j \neq j_k, k = 1,\ldots,r$, *there exists* $i(j) \in \{i_1,\ldots,i_r\}$ *such that* $(i(j), j) \in S$.

(b) There exists $j^* \neq j_k, 1 \leq k \leq r$, *such that* $(i_k, j^*) \in S$ *for all* $k = 1,\ldots,r$ *and for all* i *such that* $(i, j^*) \in S, i \neq i_k, k = 1,\ldots,r$, *there exists* $j(i) \in \{j_1,\ldots,j_r\}$ *such that* $(i, j(i)) \in S$.

Proof. It is again sufficient to show the lemma for (a). Suppose that both $\boldsymbol{X}$ and $\boldsymbol{X} + M_r(i_1,\ldots,i_r;j_1,\ldots,j_r)$ are in $\mathcal{F}(\{x_{i\cdot}\},\{x_{\cdot j}\},S)$. If this $\boldsymbol{X}$ has some positive entries at $(i^*, j_1),\ldots,(i^*, j_r)$, the decomposition in Lemma 3 can be applied and $M_r(i_1,\ldots,i_r;j_1,\ldots,j_r)$ can be replaced by r basic moves. Hence it is sufficient to consider the case of $x_{i^*j_1} + \cdots + x_{i^*j_r} = 0$. In this case, there exists some $j \neq j_k, k = 1,\ldots,r$, such that $x_{i^*j} > 0, (i^*, j) \in S$ from $x_{i^*\cdot} > 0$. We denote this j by j^*. Now from the condition (a), there exists some $i \in \{i_1,\ldots,i_r\}$ such that $(i, j^*) \in S$. We assume that i_1 satisfies this condition without loss of generality. Here, consider the following decomposition

$$
M_r(i_1,\ldots,i_r;j_1,\ldots,j_r) = M_2(i_1,i^*;j_1,j^*) + M_2(i_1,i^*;j^*,j_2) + \sum_{t=2}^{r-1} M_2(i_t,i^*;i_t,i_{t+1}) + M_2(i_r,i^*;i_r,j_1).
$$

It is again observed that the $r+1$ basic moves of the right hand side constitute a cycle, of which the consecutive two basic moves have only one nonzero cell in common. Moreover, since $x_{i^*j^*} > 0$, $\boldsymbol{X} + M_2(i_1,i^*;j^*,j_2)$ is in $\mathcal{F}(\{x_{i\cdot}\},\{x_{\cdot j}\},S)$. From these considerations, we can add these $r+1$ basic moves one by one to $\boldsymbol{X}$ without forcing negative entries on the way, instead of adding $M_r(i_1, \ldots, i_r; j_1, \ldots, j_r)$ to $\boldsymbol{X}$, and the lemma is proved. Q.E.D.

An example of the decomposition of $M_4(1,2,3,4;1,2,3,4)$ is displayed as follows.

$$
\begin{array}{|ccccc|}\hline
+1 & -1 & [0] & [0] & 0 \\
[0] & +1 & -1 & [0] & [0] \\
[0] & [0] & +1 & -1 & [0] \\
-1 & [0] & [0] & +1 & [0] \\
0 & 0 & 0 & 0 & 0 \\ \hline
\end{array}
=
\begin{array}{|ccccc|}\hline
+1 & 0 & [0] & [0] & -1 \\
[0] & 0 & 0 & [0] & [0] \\
[0] & [0] & 0 & 0 & [0] \\
0 & [0] & [0] & 0 & [0] \\
-1 & 0 & 0 & 0 & +1 \\ \hline
\end{array}
+
\begin{array}{|ccccc|}\hline
0 & -1 & [0] & [0] & +1 \\
[0] & 0 & 0 & [0] & [0] \\
[0] & [0] & 0 & 0 & [0] \\
0 & [0] & [0] & 0 & [0] \\
0 & +1 & 0 & 0 & -1 \\ \hline
\end{array}
$$

$$
+
\begin{array}{|ccccc|}\hline
0 & 0 & [0] & [0] & 0 \\
[0] & +1 & -1 & [0] & [0] \\
[0] & [0] & 0 & 0 & [0] \\
0 & [0] & [0] & 0 & [0] \\
0 & -1 & +1 & 0 & 0 \\ \hline
\end{array}
+
\begin{array}{|ccccc|}\hline
0 & 0 & [0] & [0] & 0 \\
[0] & 0 & 0 & [0] & [0] \\
[0] & [0] & +1 & -1 & [0] \\
0 & [0] & [0] & 0 & [0] \\
0 & 0 & -1 & +1 & 0 \\ \hline
\end{array}
+
\begin{array}{|ccccc|}\hline
0 & 0 & [0] & [0] & 0 \\
[0] & 0 & 0 & [0] & [0] \\
[0] & [0] & 0 & 0 & [0] \\
-1 & [0] & [0] & +1 & [0] \\
+1 & 0 & 0 & -1 & 0 \\ \hline
\end{array}
$$

Lemma 3 and 4 are concerned with replacing a particular degree r loop by a series of basic moves. We now consider the case, where all loops of degree $r \geq 3$ can be replaced by basic moves, so that a connected Markov chain over $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$ can be constructed by basic moves only.

Following Smith *et al.* (1996), consider the situation that the contingency table is $I \times I$ and the diagonal cells are structural zeros. By symmetry and as a direct consequence of Lemma 3, we see that the set of basic moves are sufficient to construct a connected chain under the assumption of positive marginals for 4×4 contingency tables with diagonal elements being structural zeros. For the case of $I \geq 5$, we see that the degree 3 loop $M_3(1,2,3;2,3,4)$ in Table 5 satisfies the condition of Lemma 4. Considering the symmetry again, we see the set of basic moves constitutes a connected Markov chain. Therefore we obtain the following corollary to Lemma 4.

Corollary 1 *A connected chain can be constructed by a set of basic moves for $I \times I$ contingency tables, $I \geq 4$, with only diagonal elements being structural zeros under the assumption of positive marginals.*

Lemma 3 and Lemma 4 give convenient sufficient conditions that a df 1 loop of degree $r \geq 3$ can be replaced by a series of basic moves. Hence for the situations that do not satisfy the conditions of Lemma 3 nor Lemma 4, we may have to use df 1 loops of degree $r \geq 3$, to ensure the connectivity of the chain. To demonstrate the importance of the minimal basis, we give an example where the set of the basic moves does not constitute a connected chain even if all the marginals are positive as follows.

$$
\begin{array}{|cccc|c}\cline{1-4}
x_{11} & x_{12} & [0] & x_{14} & 1 \\
[0] & x_{22} & x_{23} & x_{24} & 1 \\
x_{31} & [0] & x_{33} & [0] & 1 \\
[0] & x_{42} & x_{43} & x_{44} & 1 \\ \cline{1-4}
\multicolumn{1}{c}{1} & 1 & 1 & \multicolumn{1}{c}{1} &
\end{array}
$$

In this case, there are 6 elements in $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$ displayed as

$$\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, S)$$

$$= \left\{ \begin{array}{|cccc|} \hline 1 & 0 & [0] & 0 \\ [0] & 1 & 0 & 0 \\ 0 & [0] & 1 & [0] \\ [0] & 0 & 0 & 1 \\ \hline \end{array}, \begin{array}{|cccc|} \hline 1 & 0 & [0] & 0 \\ [0] & 0 & 0 & 1 \\ 0 & [0] & 1 & [0] \\ [0] & 1 & 0 & 0 \\ \hline \end{array}, \begin{array}{|cccc|} \hline 0 & 1 & [0] & 0 \\ [0] & 0 & 1 & 0 \\ 1 & [0] & 0 & [0] \\ [0] & 0 & 0 & 1 \\ \hline \end{array}, \right.$$

$$\left. \begin{array}{|cccc|} \hline 0 & 1 & [0] & 0 \\ [0] & 0 & 0 & 1 \\ 1 & [0] & 0 & [0] \\ [0] & 0 & 1 & 0 \\ \hline \end{array}, \begin{array}{|cccc|} \hline 0 & 0 & [0] & 1 \\ [0] & 1 & 0 & 0 \\ 1 & [0] & 0 & [0] \\ [0] & 0 & 1 & 0 \\ \hline \end{array}, \begin{array}{|cccc|} \hline 0 & 0 & [0] & 1 \\ [0] & 0 & 1 & 0 \\ 1 & [0] & 0 & 0 \\ [0] & 1 & 0 & 0 \\ \hline \end{array} \right\}.$$

We see that these 6 elements are not mutually reachable simply by the basic moves. To connect all the elements, we have to use degree 3 loops.

Conversely, when some of the conditions of Lemma 3 or 4 are satisfied, there is a possibility that we do not have to consider all the elements of the minimal basis. For example, if we can permute rows and columns to $\{I_1, I_2\}, \{J_1, J_2\}$, where the elements of $R(I_1; J_2), R(I_2; J_1), R(I_2; J_2)$ are all in S and there are at least one cells of S in every row and column of $R(I_1; J_1)$, we know that a set of basic moves constitutes a connected Markov chain regardless of the remaining pattern of $R(I_1; J_1)$.

8 Conclusion

As we mentioned in the introduction, a Markov chain Monte Carlo approach is extensively used in various settings of contingency tables. In our analysis of an example data displayed in Table 6, the asymptotic goodness-of-fit test overemphasizes the significance of the data and is misleading. Indeed, for many sparse tables where large sample theory does not work well, a Markov chain Monte Carlo method is a valuable tool to calculate p values for various test statistics. To construct a connected chain, a concept of the Markov basis described in this paper is essential. In this paper, we give an explicit characterization of the elements of the unique minimal Markov basis for arbitrary configurations of structural zero cells. Using the algorithm described in Section 5, we can easily obtain all the elements of the minimal basis for various problems, and we can implement a Markov chain Monte Carlo program for calculating exact p values for various test statistics. Moreover, the basis reduction described in Section 7 makes the algorithm very brief for many problems. Our experience shows that there are many problems where the set of basic moves constitutes a connected chain. For these problems, a Markov chain Monte Carlo algorithm is simply implemented by choosing pairs of rows and columns randomly.

Appendix: Minimal Markov basis for the hypothesis of quasi-symmetry

For some types of square contingency tables, the hypothesis of quasi-symmetry is of interest (Bishop *et al.*, 1975, for example). Quasi-symmetry corresponds to symmetric off-diagonal association, which implies that the interaction parameters in a log-linear model for symmetrically opposite cells are equal, but makes no assumption on the diagonal interaction parameters. Hence it is represented as $\mathrm{H}_1 : \gamma_{ij} = \gamma_{ji},\ i \neq j$ in the log-linear model (1). Quasi-symmetry model is also related to Bradley-Terry model.

For the quasi-symmetry hypothesis, an exact conditional test can also be constructed (Smith *et al.*, 1996). For this case, a sufficient statistic is $\{x_{i\cdot}\}, \{x_{\cdot j}\}, \{x_{ij} + x_{ji}\}$ and the conditional distribution is proportional to $\prod_{ij}(x_{ij}!)^{-1}$. See Smith *et al.* (1996) for detail.

To perform the exact tests of quasi-symmetry, the Markov chain Monte Carlo approach is also useful if a complete enumeration is infeasible. In this case, a connected Markov chain over the reference set

$$\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, \{x_{ij} + x_{ji}\}) = \left\{ \boldsymbol{Y} \ \middle|\ \sum_{j=1}^{J} y_{ij} = x_{i\cdot}, \sum_{i=1}^{I} y_{ij} = x_{\cdot j}, y_{ij} + y_{ji} = x_{ij} + x_{ji},\ y_{ij} \in \mathbb{N} \right\}$$

can be modified to give a connected and aperiodic Markov chain with stationary distribution as the conditional null distribution under the quasi-symmetry hypothesis by the Metropolis procedure. Similarly to the quasi-independence hypothesis, our interest is a minimal basis for a connected chain over $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, \{x_{ij} + x_{ji}\})$. The result is summarized as follows.

Definition 4 *A loop of degree r is an $I \times I$ contingency table $M_r^*(i_1, \ldots, i_r), 1 \leq i_1, \ldots, i_r \leq I$, where $M_r^*(i_1, \ldots, i_r)$ has the elements*

$$\begin{array}{l} m^*_{i_1 i_2} = m^*_{i_2 i_3} = \cdots = m^*_{i_{r-1} i_r} = m^*_{i_r i_1} = +1, \\ m^*_{i_2 i_1} = m^*_{i_3 i_2} = \cdots = m^*_{i_r i_{r-1}} = m^*_{i_1 i_r} = -1, \end{array}$$

and all the other elements are zero. Specifically, we call degree 3 loop $M_3^*(i_1, i_2, i_3)$ a basic move.

Theorem 2 *The set of the loops described in Definition 4 of degree $r = 3, \ldots, I$ constitutes a unique minimal Markov basis for $I \times I$ contingency tables under the quasi-symmetry hypothesis.*

That the set of loops described above is a Markov basis is straightforward. Uniqueness and minimality is due to the fact that $M_r^*(i_1, \ldots, i_r)$ can be written as the difference of the elements of $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, \{x_{ij} + x_{ji}\})$ which is a two-elements set.

We have the relations

$$M_r^*(i_1, \ldots, i_r) = M_r^*(i_2, i_3, \ldots, i_r, i_1) = -M_r^*(i_1, i_r, i_{r-1}, \ldots, i_2)$$

and there are

$$\sum_{r=3}^{I} \frac{I(I-1)(I-2)\cdots(I-r+1)}{2r}$$

loops in the minimal Markov basis for $I \times I$ case. Hence the algorithm can be constructed simply by generating a sequence $i_1 \cdots i_r,\ 3 \leq r \leq I$ randomly.

Similarly to Section 7, a basis reduction is also possible if all the sufficient statistics are positive.

Lemma 5 *If $x_{i\cdot}, x_{\cdot j} > 0$ for all i, j and $x_{ij} + x_{ji} > 0$ for all $i \neq j$, the set of basic moves $M_3^*(i_1, i_2, i_3)$ constitutes a connected chain.*

This lemma follows from the decomposition

$$M_r^*(i_1, \dots, i_r) = M_a^*(i_1, \dots, i_a) + M_{r-a+2}^*(i_a, i_{a+1}, \dots, i_r, i_1),$$

because the two loops of the right hand side have only two nonzero cells $m_{i_a i_1}^*, m_{i_1 i_a}^*$ in common and at least one of $x_{i_1 i_a}$ and $x_{i_a i_1}$ is positive for every element in $\mathcal{F}(\{x_{i\cdot}\}, \{x_{\cdot j}\}, \{x_{ij} + x_{ji}\})$ since $x_{ij} + x_{ji} > 0$.

References

[1] Agresti, A. (1992). A survey of exact inference for contingency tables (with discussion). *Statistical Science*, **7**, pp. 131-177.

[2] Aoki, S. and Takemura, A. (2002). Minimal basis for connected Markov chain over $3\times3\times K$ contingency tables with fixed two-dimensional marginals. Technical Report METR 02-02, Department of Mathematical Engineering and Information Physics, The University of Tokyo. Submitted for publication.

[3] Bishop, Y. M. M. and Fienberg, S. E. (1969). Incomplete two-dimensional contingency tables. *Biometrics*, **25**, pp. 119-128.

[4] Bishop, Y. M. M., Fienberg, S. E. and Holland P. W. (1975). *Discrete Multivariate Analysis: Theory and Practice.* The MIT Press, Cambridge.

[5] Besag, J. and Clifford, P. (1989). Generalized Monte Carlo significance tests. *Biometrika*, **76**, pp. 633-642.

[6] Das, T. (1945). *The Purums: An Old Kuki Tribe of Manipur.* Calcutta, University of Calcutta.

[7] Diaconis, P. and Saloff-Coste, L. (1995). Random walk on contingency tables with fixed row and column sums, Technical Report, Department of Mathematics, Harvard University.

[8] Diaconis, P. and Sturmfels, B. (1998). Algebraic algorithms for sampling from conditional distributions. *The Annals of Statistics*, **26**, pp. 363–397.

[9] Dyer, M. and Greenhill, C. (2000). Polynomial-time counting and sampling of two-rowed contingency tables, *Theoretical Computer Sciences*, **246**, pp. 265-278.

[10] Forster, J. J., McDonald, J. W. and Smith, P. W. F. (1996). Monte Carlo exact conditional tests for log-linear and logistic models, *Journal of the Royal Statistical Society, Ser. B*, **58**, pp. 445-453.

[11] Goodman, L. A. (1968). The analysis of cross-classified data: Independence, quasi-independence and interactions in contingency tables with or without missing entries. *Journal of the American Statistical Association*, **63**, pp. 1091-1131.

[12] Guo, S. W. and Thompson, E. A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrika*, **48**, pp. 361-372.

[13] Harris, J. A. (1910). On the selective elimination occurring during the development of the fruits of *Staphylea*, *Biometrika*, **7**, pp. 452-504.

[14] Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, **57**, pp. 97-109.

[15] Hernek, D. (1998). Random generation of $2 \times n$ contingency tables, *Random Structures and Algorithms*, **13**, pp. 71-79.

[16] Mantel, N. (1970). Incomplete contingency tables, *Biometrics*, **26**, pp. 291-304.

[17] Ploog, D. W. (1967). The behavior of squirrel monkeys (*Saimiri sciureus*) as revealed by sociometry, bioacoustics, and brain stimulation. In *Social Communication Among Primates*, edited by S. Altmann. pp. 149-184. Chicago. Univ. of Chicago Press.

[18] Ripley, B. D. (1987). *Stochastic Simulation*. New York, Wiley.

[19] Smith, P. W. F., Forster, J. J. and McDonald, J. W. (1996). Monte Carlo exact tests for square contingency tables, *Journal of the Royal Statistical Society, Ser A*, **159**, pp. 309-321.

[20] Takemura, A. and Aoki, S. (2002). Some characterizations of minimal Markov basis for sampling from discrete conditional distributions. Technical Report METR 02-04, Department of Mathematical Engineering and Information Physics, The University of Tokyo. Submitted for publication.

[21] Vidmar, N. (1972). Effects of decision alternatives on the verdicts and social perceptions of simulated jurors, *J. Personality and Social Psych.*, **22**, pp. 211-218.

[22] White, H. C. (1963). *An Anatomy of Kinship*, p. 138. Englewood Cliffs, N. J., Prentice-Hall.