JOURNAL OF ALGEBRAIC STATISTICS
Vol. 8, No. 1, 2017, 56-73
ISSN 1309-3452 – www.jalgstat.com

Markov bases for two-way change-point models of ladder determinantal tables

Satoshi Aoki[1,*], Takayuki Hibi[2]

[1] *Department of Mathematics, Graduate School of Science, Kobe University, Japan*
[2] *Department of Pure and Applied Mathematics, Graduate School of Information Science and Technology, Osaka University, Japan*

Abstract. To evaluate the goodness-of-fit of a statistical model to given data, calculating a conditional p value by a Markov chain Monte Carlo method is one of the effective approaches. For this purpose, a Markov basis plays an important role because it guarantees the connectivity of the chain, which is needed for unbiasedness of the estimation, and therefore is investigated in various settings such as incomplete tables or subtable sum constraints. In this paper, we consider the two-way change-point model for the ladder determinantal table, which is an extension of these two previous works, i.e., works on incomplete tables by Aoki and Takemura (2005, *J. Stat. Comput. Simulat.*) and subtable some constraints by Hara, Takemura and Yoshida (2010, *J. Pure Appl. Algebra*). Our main result is based on the theory of Gröbner basis for the distributive lattice. We give a numerical example for actual data.

2000 Mathematics Subject Classifications: 13P10, 62H17, 60J22

Key Words and Phrases: Conditional test, Contingency table, Distributive lattice, Gröbner basis, Ideal, Markov basis, Markov chain Monte Carlo, Structural zero

---

1. Introduction

In the analysis of contingency tables, computing conditional p values by a Markov chain Monte Carlo method is one of the common approaches to evaluate a fitting of a statistical model to given data. In this method, a key notion is a *Markov basis* that guarantees the connectivity of the chain for unbiasedness of the estimation. In Diaconis and Sturmfels ([7]), a notion of a Markov basis is presented with algebraic algorithms to compute it. This first work is based on a discovery of the relation between a Markov basis and a set of binomial generators of a toric ideal of a polynomial ring, which is the first connection between commutative algebra and statistics. After this first paper, Markov bases are studied intensively by many researchers both in the fields of commutative algebra and

*Corresponding author.

Email addresses: aoki@math.kobe-u.ac.jp (S. Aoki),
hibi@math.sci.osaka-u.ac.jp (T. Hibi)

statistics, which yields an attractive new field called *computational algebraic statistics*. See [15] for the first textbook of this field, and [2] for various theoretical results and examples on Markov bases.

The first result on the Markov bases in the setting of two-way contingency tables is a Markov basis for the independence model. For two-way contingency tables with fixed row sums and column sums, which is the minimal sufficient statistics under the independence model, the set of square-free moves of degree 2 forms a Markov basis. This result is generalized to the decomposable models of higher dimensional contingency tables by [8]. The reader can find various results on the structure of Markov bases of decomposable models in Chapter 8 of [2].

On the other hand, it is known that the structure of a Markov basis becomes complicated under various additional constraints to the two-way setting. One of such cases is the *incomplete two-way contingency table*, i.e., a contingency table with *structural zeros*, considered in [3]. Another case is the *subtable sum problem* considered in [9] and [14]. In these works, it is shown that moves of higher degrees are needed for Markov bases. The problem we consider in this paper is two-way contingency tables with both structural zeros and subtable sum constraints.

We consider the two-way contingency tables with specific types of structural zeros called *ladder determinantal tables*, with specific types of subtable sums called *two-way change-point model*. The two-way change-point model is considered in [13] for exponential families, including the Poisson distribution for complete two-way contingency tables. We also consider the Poisson distribution and two-way change-point model for incomplete cases in this paper. The purpose of this paper is to show that a Markov basis for this setting is constructed as the set of square-free degree 2 moves.

This paper is organized as follows. In Section 2, we illustrate the Markov chain Monte Carlo methods for the subtable sum problem of incomplete two-way contingency tables and the two-way change-point models of ladder determinant tables. In Section 3, we give the structure of the minimal Markov bases for our problems, which is the main result of this paper. The arguments and the proof of our main theorem are based on the theory of Gröbner bases for distributive lattices, which is summarized in Section 3. A numerical example for actual data is given in Section 4.

2. Preliminaries

2.1. Markov chain Monte Carlo methods for subtable sum problem of incomplete contingency tables

First we illustrate the Markov chain Monte Carlo methods for the subtable sum problem of incomplete two-way contingency tables. Though we only consider the two-way change-point model in this paper, we describe the methods in the setting of general subtable sum problems considered in [9]. Note that a specification of the subtable reduces to the two-way change-point model.

Let $\mathbb{N} = \{0, 1, 2, \ldots\}$ be the set of nonnegative integers. To consider $I \times J$ contingency tables with structural zeros, let $S \subset \{(i,j) : 1 \leq i \leq I, 1 \leq j \leq J\}$ be the set of cells that are not structural zeros. Let $q = |S|$ be the number of the cells. Let $\mathbf{x} = \{x_{ij}\} \in \mathbb{N}^q$ be an incomplete contingency table with the set of cells S, where $x_{ij} \in \mathbb{N}$ is an entry of the cell $(i,j) \in S$. Similarly to the ordinary (i.e., complete) two-way contingency tables, denote the row sums and column sums of $\mathbf{x}$ by

$$x_{i+} = \sum_{\{j\,:\,(i,j)\in S\}} x_{ij}, \quad i = 1,\ldots,I,$$

$$x_{+j} = \sum_{\{i\,:\,(i,j)\in S\}} x_{ij}, \quad j = 1,\ldots,J.$$

We assume that there is at least one $(i,j) \in S$ in each row and each column. Let B be a subset of S. We also define the subtable sum x_B by

$$x_B = \sum_{(i,j)\in B} x_{ij}.$$

Denote the set of the row sums, column sums and the subtable sum x_B by an $(I+J+1)$-dimensional column vector

$$\mathbf{t} = (x_{1+},\ldots,x_{I+},x_{+1},\ldots,x_{+J},x_B)' \in \mathbb{N}^{I+J+1}, \tag{1}$$

where $'$ is the transpose. We also treat $\mathbf{x}$ as a q-dimensional column vector as $\mathbf{x} = (x_{11}, x_{12}, \ldots, x_{IJ})'$, by lexicographic ordering of the cells in S. Then the relation between $\mathbf{x}$ and $\mathbf{t}$ is written by

$$A\mathbf{x} = \mathbf{t}, \tag{2}$$

where A is an $(I+J+1) \times p$ matrix consisting of 0's and 1's. We call A a *configuration matrix*. Though we specify S and B in Section 2.2, we show an example here.

Example 1. *Consider a* 4×4 *incomplete contingency table with* 6 *structural zeros as follows.*

x_{11}	x_{12}	x_{13}	[0]
[0]	x_{22}	x_{23}	x_{24}
[0]	[0]	x_{33}	x_{34}
[0]	[0]	x_{43}	x_{44}

In this paper, we denote a structural zero as [0] *to distinguish it from a sample zero described as* 0. *Then the set* S *is*

$$S = \{(1,1),(1,2),(1,3),(2,2),(2,3),(2,4),(3,3),(3,4),(4,3),(4,4)\}$$

and $p = 10$. *Suppose a subset* $B \subset S$ *is given by*

$$B = \{(1,1),(1,2),(1,3),(2,2),(2,3)\}.$$

Then the configuration matrix is the following 9×10 *matrix.*

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

As we see in Section 2.2, the configuration matrix considered in this paper satisfies the homogeneity assumption, i.e., the row vector $(1, \dots, 1)$ *is in the real vector space spanned by the rows of* A*. This is a natural assumption for statistical models. See Lemma 4.14 of [16] for the algebraic aspect of the homogeneity.* □

To clarify the statistical meaning of the configuration matrix A and the relation (2), consider the cell probability $\mathbf{p} = \{p_{ij}\} \in \Delta_{q-1}$, where

$$\Delta_{q-1} = \left\{ \{p_{ij}\} \in \mathbb{R}^q_{\geq 0} \,:\, \sum_{(i,j) \in S} p_{ij} = 1 \right\}$$

is called a $(q-1)$-dimensional probability simplex, and $\mathbb{R}_{\geq 0}$ is the set of nonnegative real numbers. The probability simplex Δ_{q-1} is a statistical model called a saturated model. In statistical data analysis, our interest is in a statistical model that is a subset of Δ_{q-1}. The two-way change-point model we consider in this paper is written in general form by

$$\mathcal{M} = \{\mathbf{p} = (p_{ij}) \in \Delta_{q-1} \,:\, \log p_{ij} = \alpha_i + \beta_j + \gamma \mathbf{1}_B(i,j) \text{ for some } (\alpha_i), (\beta_j), \gamma\}, \quad (3)$$

where $\mathbf{1}_B(i,j)$ is an indicator function given by

$$\mathbf{1}_B(i,j) = \begin{cases} 1, & (i,j) \in B \\ 0, & (i,j) \in S \setminus B. \end{cases}$$

Here the term $\gamma \mathbf{1}_B(i,j)$ represents a departure from the independence structure of the log-linear model. The model $\mathcal{M}$ becomes a quasi-independence model for the cells S by $\gamma = 0$. The quasi-independence model is a fundamental statistical model for the incomplete contingency tables (see Chapter 5 of [5] for detail). Sometimes, the term "quasi-independence" is also used for the model of independence except for the diagonal cells. In this paper, we use the term "quasi-independence" for a larger class of models. Markov bases for the quasi-independence model are considered in [3]. Also, the model $\mathcal{M}$ for the case that there are no structural zeros, i.e., $S = \{1, \dots, I\} \times \{1, \dots, J\}$, corresponds to the setting considered in [9]. The two-way change-point model we consider corresponds to the case

$$B = \{(i,j) \in S \,:\, i \leq i^*, j \leq j^*\} \quad (4)$$

for a fixed $(i^*, j^*) \in S$.

In this paper, we consider the fitting of the model $\mathcal{M}$ by the statistical hypothesis test

$$\begin{array}{l} \mathrm{H}_0: \ \mathbf{p} \in \mathcal{M}, \\ \mathrm{H}_1: \ \mathbf{p} \in \Delta_{p-1}. \end{array} \tag{5}$$

Under the null hypothesis H_0, $(\alpha_i), (\beta_j), \gamma$ in (3) are nuisance parameters. 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 minimal sufficient statistics for the nuisance parameters. This approach is also known as the Rao-Blackwellization of the test statistics. Using this conditional distribution, the conditional p value is defined. See [1] or Chapter 1 of [2] for detail. For our case, the minimal sufficient statistics under the null model (3) is $\mathbf{t} = A\mathbf{x}$ in (1), that is the statistical meaning of the configuration matrix A. Therefore the conditional distribution under H_0, called a *null distribution*, is written by

$$f(\mathbf{x} \mid A\mathbf{x} = \mathbf{t}) = C^{-1} \prod_{(i,j)\in S} \frac{1}{x_{ij}!},$$

where C is the normalizing constant written by

$$C = \sum_{\mathbf{y}\in\mathcal{F}_{\mathbf{t}}} \left(\prod_{(i,j)\in S} \frac{1}{y_{ij}!} \right),$$

where

$$\mathcal{F}_{\mathbf{t}} = \{\mathbf{y} \in \mathbb{N}^q \ : \ A\mathbf{y} = \mathbf{t}\}.$$

$\mathcal{F}_{\mathbf{t}}$, called a $\mathbf{t}$-*fiber*, is the set of contingency tables with given values of row sums, column sums and subtable sum. For the observed contingency table $\mathbf{x}^o$, the conditional p value for the test (5) based on a test statistic $T(\mathbf{x})$ is defined by

$$p = \sum_{\mathbf{x}\in\mathcal{F}_{A\mathbf{x}^o}} \phi(\mathbf{x}) f(\mathbf{x} \mid A\mathbf{x} = A\mathbf{x}^o),$$

where $\phi(\mathbf{x})$ is the test function of $T(\mathbf{x})$ given by

$$\phi(\mathbf{x}) = \left\{ \begin{array}{ll} 1, & T(\mathbf{x}) \geq T(\mathbf{x}^o), \\ 0, & \text{otherwise.} \end{array} \right.$$

To evaluate the conditional p value, a Monte Carlo approach is to generate samples from the null distribution $f(\mathbf{x} \mid A\mathbf{x} = A\mathbf{x}^o)$ and calculate the null distribution of the test statistics. In particular, if a connected Markov chain over $\mathcal{F}_{A\mathbf{x}^o}$ is constructed, the chain can be modified to give a connected and aperiodic Markov chain with stationary distribution $f(\mathbf{x} \mid A\mathbf{x} = A\mathbf{x}^o)$ by a Metropolis procedure, and we can use the transitions $\mathbf{x}^{(M+1)}, \mathbf{x}^{(M+2)}, \ldots \in \mathcal{F}_{A\mathbf{x}^o}$ of the chain after a large number of steps M, called burn-in steps, as samples from the null distribution. This is a *Markov chain Monte Carlo method.* See Chapter 2 of [2] or [10] for detail.

To construct a connected Markov chain over $\mathcal{F}_{A\mathbf{x}^o}$, one of the common approaches is to use a *Markov basis* introduced in [7]. An integer array $\mathbf{z} \in \mathbb{Z}^p$ satisfying $A\mathbf{z} = \mathbf{0}$ is called a *move* for the configuration A, where $\mathbb{Z}$ is the set of integers. Let

$$\mathcal{F}_0(A) = \{\mathbf{z} \in \mathbb{Z}^p \ : \ A\mathbf{z} = \mathbf{0}\}$$

denote the set of moves for A.

Definition 1 ([7])**.** *A Markov basis for A is a finite set of moves $\mathcal{B} = \{\mathbf{z}_1, \dots, \mathbf{z}_L\} \subset \mathcal{F}_0(A)$ such that, for any $\mathbf{t} \in \mathbb{N}^{I+J+1}$ and $\mathbf{x}, \mathbf{y} \in \mathcal{F}_{\mathbf{t}}$, there exist $N > 0, (\varepsilon_1, \mathbf{z}_{\ell_1}), \dots, (\varepsilon_N, \mathbf{z}_{\ell_N}) \in \mathcal{B}$ with $\varepsilon_n \in \{-1, 1\}$ such that*

$$\mathbf{y} = \mathbf{x} + \sum_{s=1}^{N} \varepsilon_s \mathbf{z}_{\ell_s} \quad \textit{and} \quad \mathbf{x} + \sum_{s=1}^{n} \varepsilon_s \mathbf{z}_{\ell_s} \in \mathcal{F}_A \ \textit{for} \ 1 \le n \le N.$$

We also define the minimality and uniqueness of the Markov basis.

Definition 2. *A Markov basis $\mathcal{B}$ is minimal if no proper subset of $\mathcal{B}$ is a Markov basis. A minimal Markov basis is unique if all minimal Markov bases differ only by sign changes of the elements.*

The fundamental results on uniqueness and minimality of Markov bases are given in Chapter 5 of [2]. For the independence model of the complete $I \times J$ contingency tables, where the minimal sufficient statistics $A\mathbf{x}$ is the row sums and column sums, it is known that the set of square-free moves of degree 2,

$$\mathcal{B} = \{\mathbf{z}(i_1, i_2; j_1, j_2), \quad 1 \le i_1 < i_2 \le I, 1 \le j_1 < j_2 \le J\},$$

where $\mathbf{z}(i_1, i_2; j_1, j_2) = \{z_{ij}\} \in \mathcal{F}_0(A)$ is given by

$$z_{ij} = \begin{cases} 1, & (i,j) = (i_1, j_1), (i_2, j_2), \\ -1, & (i,j) = (i_1, j_2), (i_2, j_1), \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

is a unique minimal Markov basis. The square-free moves of degree 2 above, displayed as

	j_1	j_2
i_1	1	-1
i_2	-1	1

,

is called a *basic move*. In the presence of the structural zeros, the set of the basic moves is not a Markov basis in general. For example, as shown in [3], incomplete tables with structural zeros as the diagonal elements, moves of degree 3 displayed as

[0]	+1	-1
-1	[0]	+1
+1	-1	[0]

are needed for Markov bases. Also, as shown in [9], if the subtable sum x_B is fixed for the patterns such as

$$(i_1, j_1), (i_2, j_2) \in B, \quad (i_1, j_2), (i_1, j_3), (i_2, j_1), (i_2, j_3) \notin B,$$

moves such as

	j_1	j_2	j_3
i_1	$+1$	$+1$	-2
i_2	-1	-1	$+2$

are needed for Markov bases. In this paper, we consider a pattern of structural zeros S, called a ladder determinantal table, and a subtable pattern (4) corresponding to a two-way change-point model and show that the set of basic moves forms a unique minimal Markov basis for this setting.

2.2. Two-way change-point models of ladder determinantal tables

Now we specify S considered in this paper.

Definition 3. *A ladder determinantal table is an incomplete contingency table with the set of cells $S \subset \{1, \ldots, I\} \times \{1, \ldots, J\}$ satisfying*

$$(1, 1), (I, J) \in S$$

and has the form

$$S = \bigcup_{i=1}^{I} \{(i, j),\ \ell_i \leq j \leq u_i\}, \tag{7}$$

where $\ell_i \leq \ell_{i+1}, u_i \leq u_{i+1}$ and $u_i \geq \ell_{i+1}$ hold for $i = 1, \ldots, I-1$.

Clearly the condition (7) is also written by

$$S = \bigcup_{j=1}^{J} \{(i, j),\ \ell'_j \leq i \leq u'_j\},$$

where $\ell'_j \leq \ell'_{j+1}, u'_j \leq u'_{j+1}$ and $u'_j \geq \ell'_{j+1}$ hold for $j = 1, \ldots, J-1$. Figure 1 illustrates examples of incomplete contingency tables. Figure 1(a) and (b) are examples of the ladder determinantal tables, whereas (c) is not. Figure 1(c) does not satisfy the condition $u_3 \geq \ell_4$ of Definition 3 because $u_3 = 3 < 4 = \ell_4$.

Remark 1. *The ladder determinantal table above is a special case of a block-stairway incomplete table. As we see in Chapter 5 of [5], an incomplete table is called a block-stairway table if it is a ladder determinantal table after permutation of rows and columns. In this paper, we do not consider permutations of rows and columns because we consider ordered categorical tables. The terminology "ladder determinantal" is used in algebraic fields. We see the relation between ladder determinantal tables and distributive lattices in Section 3.*

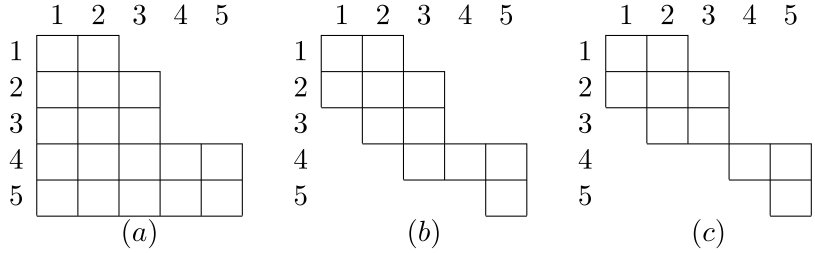

Figure 1: Examples of incomplete contingency tables. (a) and (b) are ladder determinantal tables, whereas (c) is not.

Remark 2. *The condition $u_i \geq \ell_{i+1}$ for $i = 1, \ldots, I-1$ in Definition 3 corresponds to the inseparability of incomplete tables. See Chapter 5 of [5]. We leave this condition because the inseparability is also a natural condition in our change-point models. However, it is not essential condition in our result, i.e., Theorem 2 also holds for separable incomplete tables.*

For the ladder determinantal tables $\mathbf{x}$, we consider a two-way change-point model, i.e., the model (3) with a subtable B of the form (4). Though the two-way change-point model is considered in [13] for complete contingency tables, it can be also considered for incomplete cases. We see an example in Section 4.

3. Markov bases of two-way change-point models for ladder determinantal tables

In this section we show the minimal Markov basis for two-way change-point models for ladder determinantal tables and its uniqueness. Note that the set of the basic moves, i.e, square-free moves of degree 2, is written by

$$\mathcal{B}^* = \left\{ \mathbf{z}(i_1, i_2; j_1, j_2) \;\middle|\; \begin{array}{ll} & (i_1, j_1), (i_1, j_2), (i_2, j_1), (i_2, j_2) \in B \\ \text{or} & (i_1, j_1), (i_1, j_2) \in B, (i_2, j_1), (i_2, j_2) \in S \setminus B \\ \text{or} & (i_1, j_1), (i_2, j_1) \in B, (i_1, j_2), (i_2, j_2) \in S \setminus B \\ \text{or} & (i_1, j_1), (i_1, j_2), (i_2, j_1), (i_2, j_2) \in S \setminus B \end{array} \right\},$$

where $\mathbf{z}(i_1, i_2; j_1, j_2) \in \mathcal{F}_0(A)$ is given by (6). To show the set $\mathcal{B}^*$ constitutes a Markov basis, we use the arguments of distributive lattice.

Recall that a *partial order* on a set P is a binary relation $\leq$ on P such that, for all $a, b, c \in P$, one has

- $a \leq a$ (reflexivity);
- $a \leq b$ and $b \leq a \Rightarrow a = b$ (antisymmetric);
- $a \leq b$ and $b \leq c \Rightarrow a \leq c$ (transitivity).

A partially ordered set ("poset" for short) is a set P with a partial order $\leq$. When P is a finite set, we call P a finite poset. A *lattice* is a poset L for which any two elements a and

b belonging to L possess a greatest lower bound ("meet") $a \wedge b$ and a least upper bound ("join") $a \vee b$.

Example 2. *Let B_n denote the set of subsets of $[n] = \{1, 2, \ldots, n\}$ and define the partial order $\leq$ on B_n by setting $X \leq Y$ if $X \subset Y (\subset [n])$. Then, in B_n, one has $X \cap Y = X \wedge Y$ and $X \cup Y = X \vee Y$. Thus B_n is a finite lattice, which is called the* boolean lattice *of rank n.*

A lattice L is called *distributive* if, for all $a, b, c \in L$ one has

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c), \quad a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c).$$

For example, the boolean lattice of rank n is a distributive lattice.

Let P be a finite poset. A *poset ideal* of P is a subset $\alpha \subset P$ such that

$$a \in \alpha, \ b \in P, \ b \leq a \ \Rightarrow \ b \in \alpha.$$

In particular P itself as well as the empty set $\emptyset$ is a poset ideal of P. Furthermore, if α and β are poset ideals of P, then both $\alpha \cap \beta$ and $\alpha \cup \beta$ are poset ideals of P.

Given a finite poset P, we write $L = \mathcal{J}(P)$ for the set of all poset ideals of P. We then define a partial order $\leq$ on L by setting $\alpha \leq \beta$ if $\alpha \subset \beta$, where α and β are poset ideals of P. It follows that $L = \mathcal{J}(P)$ is a finite distributive lattice.

A *totally ordered subset* of a finite poset P is a subset C of P such that, for $a, b \in C$, one has either $a \leq b$ or $b \leq a$. A totally ordered subset of P is also called a *chain* of P.

Now, a finite distributive lattice $L = \mathcal{J}(P)$ is called *planar* if

(i) P itself is *not* a chain of P;

(ii) P can be decomposed into the disjoint union of two chains of P.

Example 3. *Let $P = \{a, b, c, d\}$ be a finite poset with $a < c, b < c, b < d$. Then P is the disjoint union of chains $C = \{a, c\}$ and $D = \{b, d\}$. The finite planar distributive lattice $L = \mathcal{J}(P)$ is Figure 2.*

Suppose that $L = \mathcal{J}(P)$ is a planar distributive lattice for which P is the disjoint union of chains $C = \{a_1, \ldots, a_n\}$ and $D = \{b_1, \ldots, b_m\}$ of P with $a_1 < \cdots < a_n$ and $b_1 < \cdots < b_m$, where $n \geq 1$ and $m \geq 1$. Let

$$K[\mathbf{x}, \mathbf{y}, s, t] = K[x_1, \ldots, x_n, y_1, \ldots, y_m, s, t]$$

denote the polynomial ring in $n+m+2$ variables over a field K. We fix a poset ideal S of L with $S \neq \emptyset$ and $S \neq L$. Given $\alpha \in L$ with $i_0 = \max\{i : a_i \in \alpha\}$ and $j_0 = \max\{j : b_j \in \alpha\}$, one can associate the monomial $u_\alpha \in K[\mathbf{x}, \mathbf{y}, s, t]$ with

$$u_\alpha = \begin{cases} x_{i_0} y_{j_0} s & \text{if } \alpha \in S, \\ x_{i_0} y_{j_0} t & \text{if } \alpha \in L \setminus S. \end{cases}$$

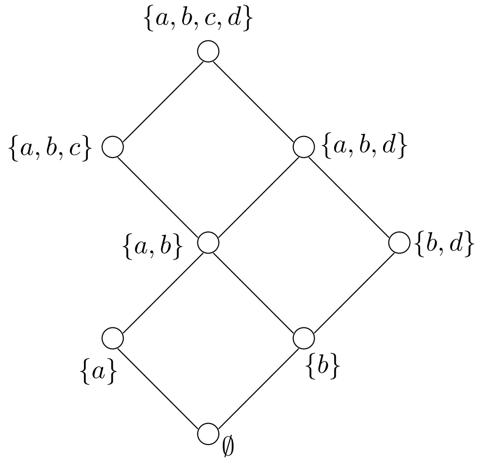

Figure 2: Distributive lattice $L = \mathcal{J}(P)$

We write $\mathcal{R}_K[L;S]$ ($\subset K[\mathbf{x},\mathbf{y},s,t]$) for the toric ring generated by those monomials u_α with $\alpha \in L$.

Let $K[L] = K[z_\alpha : \alpha \in L]$ denote the polynomial ring in $|L|$ variables over K and fix the reverse lexicographic order $<_{\text{rev}}$ on $K[L]$ induced by an ordering of the variables of $K[L]$ with the property that $z_\alpha < z_\beta$ if $\alpha < \beta$ in L. We define the surjective ring homomorphism $\pi : K[L] \to \mathcal{R}_K[L;S]$ by setting $\pi(z_\alpha) = u_\alpha$ with $\alpha \in L$. Let $I_{(L;S)}$ ($\subset K[L]$) denote the kernel of π, which will be called the *toric ideal* of $\mathcal{R}_K[L;S]$. We refer the reader to, e.g., [12] for the foundation of Gröbner bases and toric ideals.

Let $\mathcal{A}$ be the set of those 2-element subsets $\{\alpha,\beta\}$ of L, where α and β are incomparable in L, satisfying one of the following:

- $\{\alpha,\beta,\alpha\vee\beta\} \subset S$;
- $\{\alpha,\beta,\alpha\wedge\beta\} \subset L \setminus S$;
- $\alpha \in S$ and $\beta \in L \setminus S$.

It then follows that, for each $\{\alpha,\beta\} \in \mathcal{A}$, the binomial

$$f_{\alpha,\beta} = z_\alpha z_\beta - z_{\alpha\wedge\beta} z_{\alpha\vee\beta} \tag{8}$$

belongs to $I_{(L;S)}$.

Example 4. *Consider the distributive lattice for Table 1 we will consider in Section 4.*

The set of the cells of Table 1 displayed as follows.

	1	2	3	4	5	6	7
1	$(1,1)$						
2	$(2,1)$	$(2,2)$					
3	$(3,1)$	$(3,2)$	$(3,3)$				
4	$(4,1)$	$(4,2)$	$(4,3)$	$(4,4)$			
5		$(5,2)$	$(5,3)$	$(5,4)$	$(5,5)$		
6			$(6,3)$	$(6,4)$	$(6,5)$	$(6,6)$	
7				$(7,4)$	$(7,5)$	$(7,6)$	$(7,7)$

Hereafter we ignore the cells $(1,1)$ *and* $(7,7)$ *because the frequencies* x_{11} *and* x_{77} *are fixed under the model. Then the corresponding planar distributive lattice* L *is displayed in Figure 3(a). In Figure 3(a), the set of black vertices* $\bullet$ *represents a corresponding poset* P *where* $L = \mathcal{J}(P)$*, which is also displayed in Figure 3(b). Note that each vertex* $\circ$ *or* $\bullet$ *in Figure 3(a) represents a poset ideal of the poset consisting of all* $\bullet$*'s under or equal to it. For example, the vertex* $\circ$ *at* $(5,4)$ *in Figure 3(a) represents a poset ideal*

$$\{(2,2),(3,1),(3,3),(4,1),(4,4),(5,2)\},$$

of P*. The poset ideal* $S \subset L$ *of Figure 3(c) corresponds to the two-way change-point model we have considered in Section 4.*

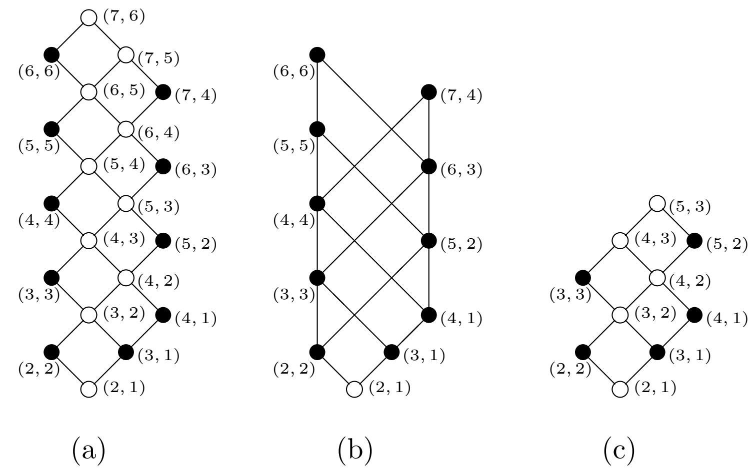

Figure 3: The planar distributive lattice for Table 1 (a), the corresponding poset (b) and the poset ideal for the two-way change-point model (c).

The poset P *is written by the disjoint union of chains*

$$C = \{(3,1),(4,1),(5,2),(6,3),(7,4)\} = \{a_3, a_4, a_5, a_6, a_7\}$$

and

$$D = \{(2,2),(3,3),(4,4),(5,5),(6,6)\} = \{b_2, b_3, b_4, b_5, b_6\}.$$

Note here that we are shifting the indices of $\{a_i\}, \{b_j\}$*, so as to correspond* a_i *to* i*-th row, and* b_j *to* j*-th column, respectively. Then for* $(i,j) \in L$*, we see that* $i_0 = i$ *and* $j_0 = j$*, and the ring homomorphism* π *is written by* $\pi(z_{ij}) = x_i y_j s$ *for* $(i,j) \in S$ *and* $\pi(z_{ij}) = x_i y_j t$ *for* $(i,j) \in L \setminus S$*, respectively.*

For the planar distributive lattice L *displayed in Figure 3(a) and for the poset ideal* $S \subset L$ *displayed in Figure 3(c), there are* 14 *incomparable* 2*-element subsets in the set* $\mathcal{A}$ *as follows.*

- $\{\alpha, \beta, \alpha \vee \beta\} \subset S$;

$$\{(2,2),(3,1)\},\{(2,2),(4,1)\},\{(3,2),(4,1)\},\{(3,3),(4,1)\},\{(3,3),(4,2)\},$$
$$\{(3,3),(5,2)\},\{(4,3),(5,2)\},$$

- $\{\alpha, \beta, \alpha \wedge \beta\} \subset L \setminus S$;

$$\{(5,5),(6,4)\},\{(5,5),(7,4)\},\{(6,5),(7,4)\},\{(6,6),(7,4)\},\{(6,6),(7,5)\},$$

- $\alpha \in S$ *and* $\beta \in L \setminus S$*:*

$$\{(4,4),(5,2)\},\{(4,4),(5,3)\}.$$

The set of the corresponding binomials (8) for these 14 *pairs coincides the set of* 14 *square-free degree* 2 *moves of (9).* □

Theorem 1. *Let* $\mathcal{G}$ *be the set of those binomials* $f_{\alpha,\beta}$ *with* $\{\alpha, \beta\} \in \mathcal{A}$*. Then* $\mathcal{G}$ *is the reduced Gröbner basis of* $I_{(L;S)}$ *with respect to* $<_{\text{rev}}$*.*

The proof of this theorem is in Appendix. From this theorem, we have the following result on the Markov basis for our problem.

Theorem 2. $\mathcal{B}^*$ *is an unique minimal Markov basis for* A *of two-way change-point models for ladder determinantal tables.*

The uniqueness of the minimal Markov basis is from the following known result.

Lemma 1 (Corollary 5.2 of [2])**.** *The unique minimal Markov basis exists if and only if the set of indispensable moves forms a Markov basis. In this case, the set of indispensable moves is the unique minimal Markov basis.*

(*Proof of Theorem 2.*) We show $\mathcal{B}^*$ corresponds to the reduced Gröbner basis of the corresponding toric ideal, and therefore a Markov basis, in Theorem 1. Because each element of $\mathcal{B}^*$ is an indispensable move, i.e., a difference of 2-element fiber, $\mathcal{B}^*$ is a unique minimal Markov basis from Lemma 1. □

4. Example

Table 1 is an example of the ladder determinantal tables from Table 4.4-13 of [5]. In this experiment, annuli from donor hydra was grafted to host hydra and observed for foot formation. The object of this experiment is to evaluate the influence of donor and grafted annulus positions on foot generation. The frequencies are the cases of foot formation out of 25 trials, and the row and column indicate the positions $1, \ldots, 7$ from foot (position 1) to head (position 7) of hydra. For this data, though it is more natural to consider binomial

Table 1: Basal disc regeneration in hydra from Table 4.4-13 of [5]

		Donor annulus position						
		1	2	3	4	5	6	7
Position of graft in host	1	4						
	2	4	0					
	3	19	5	1				
	4	24	15	4	5			
	5		19	18	18	8		
	6			24	21	16	5	
	7				23	22	8	1

sampling model, we assume Poisson sampling model here to illustrate our method. Then we consider the fitting of the two-way change-point model of

$$B = \{(1,1), (2,1), (2,2), (3,1), (3,2), (4,1), (4,2)\}.$$

The configuration matrix A is 15×22 matrix written by

$$A = \begin{pmatrix} 1000000000000000000000 \\ 0110000000000000000000 \\ 0001110000000000000000 \\ 0000001111000000000000 \\ 0000000000111100000000 \\ 0000000000000011110000 \\ 0000000000000000001111 \\ 1101001000000000000000 \\ 0010100100100000000000 \\ 0000010010010010000000 \\ 0000000001001001001000 \\ 0000000000000100100100 \\ 0000000000000000010010 \\ 0000000000000000000001 \\ 1111101100000000000000 \end{pmatrix}.$$

The fitted value of the two-way change-point model is displayed in Table 2. As a test

Table 2: Fitted value of the two-way change-point model $(i^*, j^*) = (4, 2)$ for Table Table 4.4-13 of 1

		Donor annulus position						
		1	2	3	4	5	6	7
	1	4.00						
	2	2.81	1.19					
Position of graft	3	15.94	6.78	2.28				
in host	4	28.26	12.03	4.05	3.67			
	5		19.00	17.17	15.54	11.29		
	6			23.50	21.27	15.45	5.79	
	7				26.52	19.26	7.21	1.00

statistic, we use Pearson's goodness-of-fit χ^2

$$\chi^2 = \sum_{(i,j)\in S} \frac{(x_{ij} - m_{ij})^2}{m_{ij}},$$

where $\mathbf{m} = (m_{ij})$ is the fitted value in Table 2. We have $\chi^2 = 7.814$ with 8 degrees of freedom. From Theorem 2, a unique minimal Markov basis is the set of 14 square-free degree 2 moves below,

$$\begin{array}{lllll} \mathbf{z}(2,3;1,2), & \mathbf{z}(2,4;1,2), & \mathbf{z}(3,4;1,2), & \mathbf{z}(3,4;1,3), & \mathbf{z}(3,4;2,3), \\ \mathbf{z}(4,5;3,4), & \mathbf{z}(4,6;3,4), & \mathbf{z}(5,6;3,4), & \mathbf{z}(5,6;3,5), & \mathbf{z}(5,6;4,5), \\ \mathbf{z}(5,7;4,5), & \mathbf{z}(6,7;4,5), & \mathbf{z}(6,7;4,6), & \mathbf{z}(6,7;5,6), & \end{array} \qquad (9)$$

where $\mathbf{z}(i_1, i_2; j_1, j_2)$ is given by (6). Using the above Markov basis, we calculate the conditional p value by the Markov chain Monte Carlo method. Starting from the observed data, after discarding 50000 burn-in samples, we generate 100000 samples from the Markov chain and have the estimate $\hat{p} = 0.46$. Note that the asymptotic p value based on the asymptotic χ^2_8 distribution of the test statistics is 0.452, which means good fitting of the asymptotic distribution for Table 1. Figure 4 is a histogram of Pearson's goodness-of-fit χ^2 generated by the Markov chain, which also shows the good fitting of the asymptotic distribution. Similarly, we check the goodness-of-fits of all the two-way change-point models for each (i^*, j^*), and find that the model with $(i^*, j^*) = (4, 2)$ is the best two-way change-point model for Table 1, i.e., the model with the maximal estimated p value.

5. Discussion

In this paper, we give a unique minimal Markov basis for two-way change-point models of ladder determinantal tables. Our setting is an extension of two papers, [3] and [9]. The

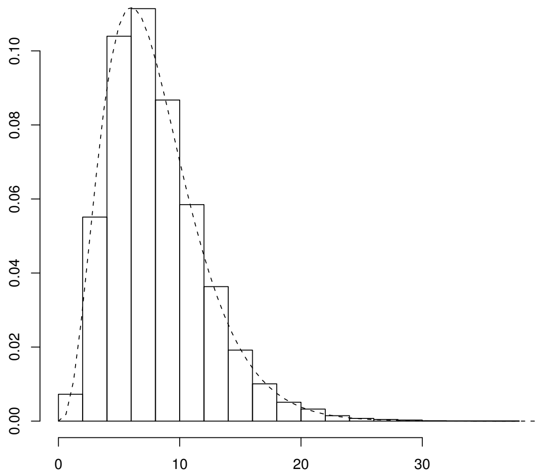

Figure 4: A histogram of Pearson's goodness-of-fit χ^2 generated by the Markov chain. Dotted line is the asymptotic χ^2_8 distribution.

two-way change-point model is an example of subtable sum problems considered in [9], and the ladder determinantal table is an example of incomplete contingency tables considered in [3]. We consider both constraints at once in this paper.

Our main result is based on the theory of Gröbner bases for the distributive lattice. As we see in Section 3, the ladder determinantal table is treated as the distributive lattice. One important point is that we can consider any poset ideal as the two-way change-point models, even if it is not a rectangular shape as (4). Therefore our method is also used for any B as long as it corresponds to a poset ideal of the distributive lattice.

In the analysis of two-way contingency tables, several extensions of the independence model are considered from the viewpoint of algebraic statistics. For example, a weakened independence model by [6] is constructed from the set of 2×2 adjacent minors.

A. Proof of Theorem 1

Proof. Once we know that $\mathcal{G}$ is a Gröbner basis of $I_{(L;S)}$ with respect to $<_{\rm rev}$, it follows immediately that $\mathcal{G}$ is reduced. The initial monomial ${\rm in}_{<_{\rm rev}}(f_{\alpha,\beta})$ of $f_{\alpha,\beta}$ is ${\rm in}_{<_{\rm rev}}(f_{\alpha,\beta}) = z_\alpha z_\beta$. Let ${\rm in}_{<_{\rm rev}}(\mathcal{G})$ denote the ideal of $K[L]$ generated by those monomials ${\rm in}_{<_{\rm rev}}(f_{\alpha,\beta})$ with $f_{\alpha,\beta} \in \mathcal{G}$. Clearly ${\rm in}_{<_{\rm rev}}(\mathcal{G}) \subset {\rm in}_{<_{\rm rev}}(I_{(L;S)})$, where ${\rm in}_{<_{\rm rev}}(I_{(L;S)})$ is the initial ideal of $I_{(L;S)}$ with respect to $<_{\rm rev}$. In order to show that $\mathcal{G}$ is a Gröbner basis of $I_{(L;S)}$ with respect to $<_{\rm rev}$, by virtue of the technique [4, Lemma 1.1], what we must prove is that, for monomials u and v, where $u \neq v$, belonging to $K[L]$ with $u \notin {\rm in}_{<_{\rm rev}}(\mathcal{G})$ and $v \notin {\rm in}_{<_{\rm rev}}(\mathcal{G})$, one has $\pi(u) \neq \pi(v)$. One can assume that u and v are relatively prime and, furthermore,

$$u = z_{\alpha_1} \cdots z_{\alpha_p} z_{\beta_1} \cdots z_{\beta_q}, \quad v = z_{\alpha'_1} \cdots z_{\alpha'_p} z_{\beta'_1} \cdots z_{\beta'_q},$$

where each $\alpha_i \in S$, each $\alpha'_i \in S$, each $\beta_j \in L \setminus S$ and each $\beta'_j \in L \setminus S$. Since $z_\alpha z_\beta \in \mathrm{in}_{<_{\mathrm{rev}}}(\mathcal{G})$ if α and β are incomparable in L with $\alpha \in S$ and $\beta \in L \setminus S$, the condition that

($\sharp$) for each i and for each j, one has $\alpha_i < \beta_j$ and $\alpha'_i < \beta'_j$

is satisfied.

If $\alpha_i \vee \alpha_{i'} \in S$, then α_i and $\alpha_{i'}$ must be comparable in L. Thus in particular, if $\alpha_i \vee \alpha_{i'} \in S$ for each i and for each i' with $1 \leq i < i' \leq p$, then $\{\alpha_1, \ldots, \alpha_p\}$ is a multichain of L. On the other hand, suppose that there exist $1 \leq i < i' \leq p$ with $\alpha_i \vee \alpha_{i'} \in L \setminus S$. Then, by ($\sharp$), for each j and for each j' with $1 \leq j < j' \leq q$, one has $\beta_j \wedge \beta_{j'} \in L \setminus S$, so that β_j and $\beta_{j'}$ must be comparable in L. Hence $\{\beta_1, \ldots, \beta_q\}$ is a multichain of L.

Now, suppose that

(i) for each i and for each i' with $1 \leq i < i' \leq p$, one has $\alpha_i \vee \alpha_{i'} \in S$;

(ii) there exist $1 \leq k < k' \leq p$ for which $\alpha'_k \vee \alpha'_{k'} \in L \setminus S$.

Then each of $\{\alpha_1, \ldots, \alpha_p\}$ and $\{\beta'_1, \ldots, \beta'_q\}$ is a multichain of L. Ignoring the variables s and t, the toric ring $\mathcal{R}_K[L]$ introduced in [11] arises. Working in the frame of [11], if $u^* = z_{\gamma_1} \cdots z_{\gamma_{p+q}}$ is the standard monomial expression of u and $v^* = z_{\gamma'_1} \cdots z_{\gamma'_{p+q}}$ is that of v, then again by ($\sharp$) one has $|\{i : \gamma_i \in S\}| \geq p$ and $|\{j : \gamma'_j \in S\}| < p$. Hence $u^* \neq v^*$. Thus $\pi(u) \neq \pi(v)$.

The same argument as above shows that if we suppose

(i') for each j and for each j' with $1 \leq j < j' \leq q$, one has $\beta_j \wedge \beta_{j'} \in L \setminus S$;

(ii') there exist $1 \leq \ell < \ell' \leq q$ for which $\beta'_\ell \wedge \beta'_{\ell'} \in S$,

then $\pi(u) \neq \pi(v)$.

Let $\pi(u) = \pi(v)$. Then one can assume one of the following conditions:

($\clubsuit$) for each i and for each i' with $1 \leq i < i' \leq p$, one has $\alpha_i \vee \alpha_{i'} \in S$ and $\alpha'_i \vee \alpha'_{i'} \in S$;

($\spadesuit$) for each j and for each j' with $1 \leq i < i' \leq q$, one has $\beta_j \wedge \beta_{j'} \in L \setminus S$ and $\beta'_j \wedge \beta'_{j'} \in L \setminus S$.

Suppose ($\clubsuit$). Then each of $\{\alpha_1, \ldots, \alpha_p\}$ and $\{\alpha'_1, \ldots, \alpha'_p\}$ is a multichain of L. Hence, by ($\sharp$) together with [11], one has $\{\alpha_1, \ldots, \alpha_p\} = \{\alpha'_1, \ldots, \alpha'_p\}$ as multichains of L. Since u and v are relatively prime, one has $p = 0$.

Let $p = 0$ and $q \geq 2$. Let $\pi(z_{\beta_j}) = x_{\xi_j} y_{\zeta_j} t$ for $1 \leq j \leq q$. Set $\xi = \min\{\xi_j : 1 \leq j \leq q\}$ and write ζ for the smallest integer for which there is $1 \leq j_0 \leq q$ with $\pi(z_{\beta_{j_0}}) = x_\xi y_\zeta t$. Then there exist β'_{j_1} and β'_{j_2} such that $\pi(\beta'_{j_1}) = x_\xi y_{j_*} t$ and $\pi(\beta'_{j_2}) = x_{i_*} y_\zeta t$. One has $i_* > \xi$ and $j_* > \zeta$. Hence $\beta'_{j_1} \wedge \beta'_{j_2} = \beta_{j_0}$. Since $\beta_{j_0} \in L \setminus S$ and since β'_{j_1} and β'_{j_2} are incomparable in L, one has $z_{\beta'_{j_1}} z_{\beta'_{j_2}} \in \mathrm{in}_{<_{\mathrm{rev}}}(\mathcal{G})$, which contradicts $v \notin \mathrm{in}_{<_{\mathrm{rev}}}(\mathcal{G})$.

Finally, the same argument as above is also valid if we suppose ($\spadesuit$).

This completes proving that $\mathcal{G}$ is the reduced Gröbner basis of $I_{(L;S)}$ with respect to $<_{\mathrm{rev}}$.

References

[1] Alan Agresti. A survey of exact inference for contingency tables. *Statistical Science*, 7:131–177, 1992.

[2] Satoshi Aoki, Hisayuki Hara, and Akimichi Takemura. *Markov bases in algebraic statistics.* Springer Series in Statistics., New York, 2012.

[3] Satoshi Aoki and Akimichi Takemura. Markov chain monte carlo exact tests for incomplete two-way contingency tables. *J. Stat. Comput. Simulat.*, 75:787–812, 2005.

[4] Annetta Aramova, Jürgen Herzog, and Takayuki Hibi. Finite lattices and lexicographic gröbner bases. *Europ. J. Combin.*, 21:431–439, 2000.

[5] Yvonne M.M. Bishop, Stephen E. Fienberg, and Paul W. Holland. *Discrete multivariate analysis, Theory and applications.* The MIT Press, Cambridge, Massachusetts, 1975.

[6] Enrico Carlini and Fabio Rapallo. A class of statistical models to weaken independence in two-way contingency tables. *Metrika*, 73:1–22, 2011.

[7] Persi Diaconis and Bernd Sturmfels. Algebraic algorithms for sampling from conditional distributions. *Annals of Statistics*, 26:363–397, 1998.

[8] Adrian Dobra. Markov bases for decomposable graphical models. *Bernoulli*, 9:1093–1108, 2003.

[9] Hisayuki Hara, Akimichi Takemura, and Ruriko Yoshida. Markov bases for subtable sum problems. *J. Pure Appl. Algebra*, 213:1507–1521, 2010.

[10] Wilfred Keith Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57:97–109, 1970.

[11] Takayuki Hibi. Distributive lattices, affine semigroup rings and algebras with straightening laws. In M. Nagata and H. Matsumura, editors, *Commutative Algebra and Combinatorics.*, volume 11, pages 93–109. Advanced Studies in Pure Math., North–Holland, Amsterdam, 1987.

[12] Takayuki (Ed.) Hibi. *Gröbner Bases: Statistics and Software Systems.* Springer, 2013.

[13] Chihiro Hirotsu. Two-way change-point model and its application. *Austral. J. Statist.*, 39:205–218, 1997.

[14] Hidefumi Ohsugi and Takayuki Hibi. Two way subtable sum problems and quadratic gröbner bases. *Proc. Amer. Math. Soc.*, 137:1539–1542, 2009.

[15] Giovanni Pistone, Eva Riccomagno, and Henry Wynn. *Algebraic statistics: Computational commutative algebra in statistics.* Chapman & Hall Ltd, Boca Raton, 2001.

[16] Bernd Sturmfels. *Gröbner bases and convex polytopes.* University Lecture Series, **8**, American Mathematical Society, Providence, RI., 1996.