

The Set of $3 \times 4 \times 4$ Contingency Tables has 3-Neighborhood Property

Toshio Sumi and Toshio Sakata

Faculty of Design, Kyushu University

COMPSTAT2010
25 August 2010
in Paris, France

- 1 Three way contingency tables
- 2 Motivation
- 3 r neighbourhood property and the main results
- 4 Markov basis
- 5 General Perspectives
- 6 $3 \times 3 \times 4$ contingency tables
- 7 Future work

- One way contingency tables $\dots\dots$ vectors (h_i)
- Two way contingency tables $\dots\dots$ matrices (h_{ij})
- Three way contingency tables $\dots\dots$ (h_{ijk})

An $I \times J \times K$ contingency table $\dots\dots$ (h_{ijk}) , where $1 \leq i \leq I$, $1 \leq j \leq J$, and $1 \leq k \leq K$. $I \times J \times K$ contingency tables are 1 – 1 corresponding with functions from $[1..I] \times [1..J] \times [1..K]$ to $\mathbb{Z}_{\geq 0}$.

$$[1..n] := \{1, 2, \dots, n\}$$

In the analysis of three-way contingency tables we often use the conditional inference and recently the conditional test of three way tables has seen some enthusiasm. (cf. Diaconis and Sturmfels: 1998, Aoki and Takemura: 2003, 2004). In $I \times J \times K$ three-way tables the probability function is given by

$$P\{X = \mathbf{x} \mid \mathbf{p}\} \sim \frac{\prod_{(i,j,k) \in \mathcal{Z}} p_{ijk}^{x_{ijk}}}{\prod_{(i,j,k) \in \mathcal{Z}} x_{ijk}!},$$

where $\mathcal{Z} = \{(i, j, k) \mid 1 \leq i \leq I, 1 \leq j \leq J, 1 \leq k \leq K\}$, $\mathbf{x} = (x_{ijk}; (i, j, k) \in \mathcal{Z})$ is a family of cell counts, and $\mathbf{p} = (p_{ijk}; (i, j, k) \in \mathcal{Z})$ is a family of cell probabilities.

In the log-linear model the probability p_{ijk} is expressed as

$$\log p_{ijk} = \mu + \tau_i^X + \tau_j^Y + \tau_k^Z + \tau_{ij}^{XY} + \tau_{jk}^{YZ} + \tau_{ik}^{XZ} + \tau_{ijk}^{XYZ}$$

where τ^{XYZ} is called three-way interaction effect (Agresti:1996).

The hypothesis to be tested is

$$H : \tau_{ijk}^{XYZ} = 0 \quad \text{for all } (i, j, k) \in \mathcal{Z}$$

which means that there is no three-way interaction. Under the hypothesis H the sufficient statistic is the set of two way x - y , y - z , x - z marginals and so the conditional probabilities becomes free from parameters under H .

Upon fixing all two way marginals the conditional distribution of X becomes

$$P_H\{X = x \mid \alpha, \beta, \gamma\} = \frac{\prod_{(i,j,k) \in \mathcal{Z}} 1/x_{ijk}!}{\sum_{y \in \mathcal{F}} \prod_{(i,j,k) \in \mathcal{Z}} 1/y_{ijk}!},$$

where $\mathcal{F} = \mathcal{F}(\alpha, \beta, \gamma)$ is the set of three-way contingency tables with the two-way marginals, and α , β and γ are the x - y , y - z , x - z marginals of the observed table respectively. The important thing is that the distribution is of parameter free under H . When $X = x_0$ was observed, our primary concern is to evaluate the probability

$$\mathbf{p}\text{-value} = P_H\{T(X) \geq T(x_0)\},$$

where T is an appropriate test statistic.

Let $\mathcal{F}(\alpha, \beta, \gamma)$ be the set of contingency tables with marginals α, β, γ and $\mathcal{F}(H)$ the set of contingency tables with marginals as same as those of H .

To evaluate the p -value, we consider about the Monte Carlo method. The Monte Carlo method estimates $\mathcal{F}(\alpha, \beta, \gamma)$ by running a Markov chain. The Markov chain must be irreducible and in order to generate an irreducible Markov chain we need a Markov basis \mathcal{B} by which all elements in $\mathcal{F}(\alpha, \beta, \gamma)$ become mutually reachable by a sequence of elements in \mathcal{B} without violating non-negativity condition.

In the $(t - 1)$ -stage, for a given dataset we obtain a contingency table H_{t-1} and let consider the set $\mathcal{F}(H_{t-1})$. If one data is obtained at (i_t, j_t, k_t) , we have a new contingency table H_t by combining it with the given dataset and consider the set $\mathcal{F}(H_t)$.

$$H_t(\cdot, j, k) = \begin{cases} H_{t-1}(\cdot, j, k) & (j, k) \neq (j_t, k_t) \\ H_{t-1}(\cdot, j_t, k_t) + 1 & (j, k) = (j_t, k_t) \end{cases}$$

$$H_t(i, \cdot, k) = \begin{cases} H_{t-1}(i, \cdot, k) & (i, k) \neq (i_t, k_t) \\ H_{t-1}(i_t, \cdot, k_t) + 1 & (i, k) = (i_t, k_t) \end{cases}$$

$$H_t(i, j, \cdot) = \begin{cases} H_{t-1}(i, j, \cdot) & (i, j) \neq (i_t, j_t) \\ H_{t-1}(i_t, j_t, \cdot) + 1 & (i, j) = (i_t, j_t). \end{cases}$$

In the sequential conditional test, consider

$$\begin{aligned} \mathcal{F}(H_1) &\rightarrow \mathcal{F}(H_2) \rightarrow \dots \\ &\rightarrow \mathcal{F}(H_{t-1}) \rightarrow \mathcal{F}(H_t) \rightarrow \dots . \end{aligned}$$

Although MCMC test by Metropolis-Hastings's algorithm is general in the sequential conditional test, our purpose is to obtain $\mathcal{F}(H_t)$ by using the previous $\mathcal{F}(H_{t-1})$ and completely exact probabilities in Fisher's exact test.

Programming by R for $3 \times 3 \times 3$ contingency tables

Step t	$ \mathcal{F}_t $	Ours	MCMC1	MCMC2
21	12	0.2727273	0.2692	0.2715
22	15	0.3628319	0.3622	0.3628
23	19	0.4824798	0.4952	0.4747
24	25	0.3872708	0.3766	0.3886
25	32	0.1602634	0.1628	0.1618
26	99	0.1176134	0.123	0.1203
27	144	0.05369225	0.0534	0.0503
28	152	0.03016754	0.0322	0.0291

MCMC1: $(5 * 10^3, 5 * 10^2)$

Select $5 * 10^3$ tables each of which is got by $5 * 10^2$ skip.

MCMC2: $(10^4, 10^3)$

Step t	$ \mathcal{F}_t $	Ours	MCMC1	MCMC2
21	12	0.016	74.464	297.655
22	15	0.033	73.699	302.268
23	19	0.04	75.482	309.072
24	25	0.046	77.406	314.035
25	32	0.089	80.639	326.231
26	99	0.232	88.107	354.706
27	144	0.275	91.275	368.7
28	152	0.138	91.899	372.771
		0.869	652.971	2645.438

MCMC1: $(5 * 10^3, 5 * 10^2)$

MCMC2: $(10^4, 10^3)$

Step t	$ \mathcal{F}_t $	Ours	MCMC1	MCMC2	Prob.
21	4	0.026	66.655	282.103	1.0
31	63	0.045	83.659	337.438	0.2169823
42	253	0.565	96.445	390.396	0.2925166
50	1168	0.132	114.436	464.128	0.4415128
51	1493	1.961	117.895	479.225	0.4047599
60	6663	15.482	141.59	572.312	0.1865068
61	11599	42.942	151.726	617.728	0.1059830
71	15784	0.556	154.862	626.537	0.06565988
72	17285	15.624	154.978	628.453	0.05635573
73	17285	0.727	154.922	626.177	0.04961687
		149.51	5921.512	23969.86	

MCMC1: $(5 * 10^3, 5 * 10^2)$

MCMC2: $(10^4, 10^3)$

Step t	$ \mathcal{F}_t $	Ours	MCMC1	MCMC2	Prob.
21	4	0.026	66.655	282.103	1.0
31	63	0.045	83.659	337.438	0.2169823
42	253	0.565	96.445	390.396	0.2925166
50	1168	0.132	114.436	464.128	0.4415128
51	1493	1.961	117.895	479.225	0.4047599
60	6663	15.482	141.59	572.312	0.1865068
61	11599	42.942	151.726	617.728	0.1059830
71	15784	0.556	154.862	626.537	0.06565988
72	17285	15.624	154.978	628.453	0.05635573
73	17285	0.727	154.922	626.177	0.04961687
		149.51	5921.512	23969.86	

MCMC1: $(5 * 10^3, 5 * 10^2)$

MCMC2: $(10^4, 10^3)$

Step t	$ \mathcal{F}_t $	Ours	MCMC1	MCMC2	Prob.
21	4	0.026	66.655	282.103	1.0
31	63	0.045	83.659	337.438	0.2169823
42	253	0.565	96.445	390.396	0.2925166
51	1493	1.961	117.895	479.225	0.4047599
60	6663	15.482	141.59	572.312	0.1865068
61	11599	42.942	151.726	617.728	0.1059830
71	15784	0.556	154.862	626.537	0.06565988
72	17285	15.624	154.978	628.453	0.05635573
73	17285	0.727	154.922	626.177	0.04961687
		149.51	5921.512	23969.86	

MCMC1: $(5 * 10^3, 5 * 10^2)$

MCMC2: $(10^4, 10^3)$

$I \times J \times K$ contingency table

$I \times J \times K$ contingency table consists of K slices of $I \times J$ matrices consisting non-negative integers.

$i \setminus j$ $k = 1$

h_{111}	h_{121}	h_{131}
h_{211}	h_{221}	h_{231}
h_{311}	h_{321}	h_{331}

$i \setminus j$ $k = 2$

h_{112}	h_{122}	h_{132}
h_{212}	h_{222}	h_{232}
h_{312}	h_{322}	h_{332}

$i \setminus j$ $k = 3$

h_{113}	h_{123}	h_{133}
h_{213}	h_{223}	h_{233}
h_{313}	h_{323}	h_{333}

Table: $3 \times 3 \times 3$ contingency table

Marginals for an $I \times J \times K$ contingency table

$i \setminus j$ x-y marginal

$h_{11\cdot}$	$h_{12\cdot}$	$h_{13\cdot}$
$h_{21\cdot}$	$h_{22\cdot}$	$h_{23\cdot}$
$h_{31\cdot}$	$h_{32\cdot}$	$h_{33\cdot}$

$j \setminus k$ y-z marginal

$h_{\cdot 11}$	$h_{\cdot 12}$	$h_{\cdot 13}$
$h_{\cdot 21}$	$h_{\cdot 22}$	$h_{\cdot 23}$
$h_{\cdot 31}$	$h_{\cdot 32}$	$h_{\cdot 33}$

$i \setminus k$ x-z marginal

$h_{1\cdot 1}$	$h_{1\cdot 2}$	$h_{1\cdot 3}$
$h_{2\cdot 1}$	$h_{2\cdot 2}$	$h_{2\cdot 3}$
$h_{3\cdot 1}$	$h_{3\cdot 2}$	$h_{3\cdot 3}$

Table: Marginals of a $3 \times 3 \times 3$ contingency table

$$h_{ij\cdot} = \sum_{s=1}^K h_{ijs}, \quad h_{\cdot jk} = \sum_{s=1}^I h_{sjk}, \quad \text{and} \quad h_{i\cdot k} = \sum_{s=1}^J h_{isk}$$

Find $\mathcal{F}(H_t)$ from $\mathcal{F}(H_{t-1})$.

Put $\mathcal{F}_t = \mathcal{F}(H_t)$ for any t .

Let ϕ_t be a map from \mathcal{F}_{t-1} to \mathcal{F}_t by simply adding 1 in the (i_t, j_t, k_t) -cell.

Remark

A table T of \mathcal{F}_t with $T_{i_j k_t} > 0$ lies in the image of ϕ_t .

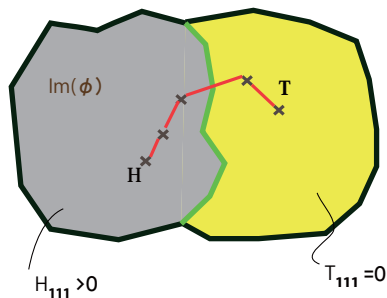
Thus we may find all tables T of \mathcal{F}_t with $T_{i_j k_t} = 0$.

From now on we assume $(i_t, j_t, k_t) = (1, 1, 1)$ for simplicity. By the above remark, we need to consider how we can generate $H \in \mathcal{F}_t$ with $H_{111} = 0$.

From now on we assume $(i_t, j_t, k_t) = (1, 1, 1)$ for simplicity. By the above remark, we need to consider how we can generate $H \in \mathcal{F}_t$ with $H_{111} = 0$.

An idea is to use a Markov basis.

From now on we assume $(i_t, j_t, k_t) = (1, 1, 1)$ for simplicity. By the above remark, we need to consider how we can generate $H \in \mathcal{F}_t$ with $H_{111} = 0$.



For any $T, H \in \mathcal{F}_t$, there is a sequence of moves F_1, \dots, F_s of the Markov basis such that

$$H_1 := H + F_1 \in \mathcal{F}_t$$

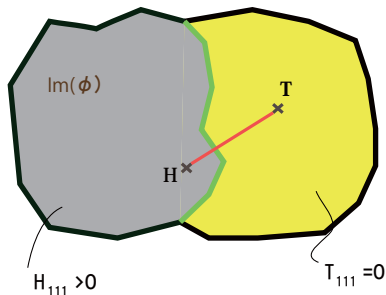
$$H_2 := H_1 + F_2 \in \mathcal{F}_t$$

$$\vdots$$

$$H_s := H_{s-1} + F_s \in \mathcal{F}_t$$

$$T = H_s$$

From now on we assume $(i_t, j_t, k_t) = (1, 1, 1)$ for simplicity. By the above remark, we need to consider how we can generate $H \in \mathcal{F}_t$ with $H_{111} = 0$.



We want to find a set of moves \mathcal{B} such that for any $T \in \mathcal{F}_t$ with $T_{111} = 0$, there are $H \in \phi(\mathcal{F}_{t-1})$ and $F \in \mathcal{B}$ such that $T + F = H$.

Markov move is a table with all zero marginal. A set \mathcal{B} of Markov moves is called a *Markov basis* if for an arbitrary two contingency tables H and H' with the same marginals, say α, β, γ , there are Markov moves M_1, \dots, M_r (for some r) in \mathcal{B} such that

$$\begin{aligned} &H + M_1, \\ &(H + M_1) + M_2, \\ &\quad \vdots \\ &(\cdots (H + M_1) + \cdots + M_{r-1}) + M_r = H' \end{aligned}$$

are all contingency tables in $\mathcal{F}(H)$. A *minimal* Markov basis has a minimality property in the set of Markov basis.

A minimal Markov basis for $3 \times 4 \times 4$ contingency tables as we use in this talk and it is unique by Aoki-Takemura.

A minimal Markov basis for $3 \times 4 \times 4$ contingency tables as we use in this talk and it is unique by Aoki-Takemura.

We fix a **minimal** Markov basis \mathcal{B} . For H and $H' \in \mathcal{F}_t$, H' is said to be in the r -neighbourhood of H if H' is reachable from H by at most r moves of \mathcal{B} . \mathcal{F}_t has r -neighbourhood property if for each $H \in \mathcal{F}_t$ there is $H' \in \mathcal{F}_t$ with $H'_{111} > 0$ in the r -neighbourhood of H and there is $H \in \mathcal{F}_t$ such that the $(r - 1)$ -neighbourhood of H has no H' with $H'_{111} > 0$.

Markov basis for $3 \times 4 \times 4$ contingency tables

Aoki and Takemura determined a minimal Markov basis and showed it is unique.

Theorem (Aoki-Takemura)

The set of $222_4(i_1 i_2, j_1 j_2, k_1 k_2)$, $233_6(i_1 i_2, j_1 j_2 j_3, k_1 k_2 k_3)$, $323_6(i_1 i_2 i_3, j_1 j_2, k_1 k_2 k_3)$, $332_6(i_1 i_2 i_3, j_1 j_2 j_3, k_1 k_2)$, $244_8(i_1 i_2, j_1 j_2 j_3 j_4, k_1 k_2 k_3 k_4)$, $334_8(i_1 i_2 i_3, j_1 j_2 j_3, k_1 k_2 k_3 k_4)$, $344_9(i_1 i_2 i_3, j_1 j_2 j_3 j_4, k_1 k_2 k_3 k_4)$, and $344_{10}(i_1 i_2 i_3, j_1 j_2 j_3 j_4, k_1 k_2 k_3 k_4)$ is a minimal basis for $3 \times 4 \times 4$ contingency tables.

▶ skip Markov basis

Markov move of degree 6

$233_6(i_1 i_2, j_1 j_2 j_3, k_1 k_2 k_3)$ is a move of degree 6 so that the cells of (i_1, j_1, k_1) , (i_1, j_2, k_2) , (i_1, j_3, k_3) , (i_2, j_1, k_2) , (i_2, j_2, k_3) , and (i_2, j_3, k_1) take 1, the cells of (i_1, j_1, k_2) , (i_1, j_2, k_3) , (i_1, j_3, k_1) , (i_2, j_1, k_1) , (i_2, j_2, k_2) , and (i_2, j_3, k_3) take -1 , and all the other cells are zero.

$i \setminus j$	$k = k_1$	$k = k_2$	$k = k_3$
	$j = j_1$	$j = j_2$	$j = j_3$
$i = i_1$	1		-1
$i = i_2$	-1		1

$i \setminus j$	$k = k_1$	$k = k_2$	$k = k_3$
	$j = j_1$	$j = j_2$	$j = j_3$
$i = i_1$	-1	1	
$i = i_2$	1	-1	

$i \setminus j$	$k = k_1$	$k = k_2$	$k = k_3$
	$j = j_1$	$j = j_2$	$j = j_3$
$i = i_1$		-1	1
$i = i_2$		1	-1

Markov move of degree 6

$233_6(i_1 i_2, j_1 j_2 j_3, k_1 k_2 k_3)$ is a move of degree 6 so that the cells of (i_1, j_1, k_1) , (i_1, j_2, k_2) , (i_1, j_3, k_3) , (i_2, j_1, k_2) , (i_2, j_2, k_3) , and (i_2, j_3, k_1) take 1, the cells of (i_1, j_1, k_2) , (i_1, j_2, k_3) , (i_1, j_3, k_1) , (i_2, j_1, k_1) , (i_2, j_2, k_2) , and (i_2, j_3, k_3) take -1 , and all the other cells are zero.

$i \setminus j$	$j = j_1$	$j = j_2$	$j = j_3$	$j = j_1$	$j = j_2$	$j = j_3$	$j = j_1$	$j = j_2$	$j = j_3$
$k = k_1$									
$i = i_1$	1		-1	-1	1		-1	-1	1
$i = i_2$	-1		1	1	-1		1	1	-1

$$233_6(i_1 i_2, j_1 j_2 j_3, k_1 k_2 k_3) = 222_4(i_1 i_2, j_1 j_3, k_1 k_3) + 222_4(i_2 i_1, j_1 j_2, k_2 k_3)$$

Markov move of degree 6

$323_6(i_1 i_2 i_3, j_1 j_2, k_1 k_2 k_3)$ is a move of degree 6 so that the cells of (i_1, j_1, k_1) , (i_2, j_1, k_2) , (i_3, j_1, k_3) , (i_1, j_2, k_2) , (i_2, j_2, k_3) , and (i_3, j_2, k_1) take 1, the cells of (i_1, j_1, k_2) , (i_2, j_1, k_3) , (i_3, j_1, k_1) , (i_1, j_2, k_1) , (i_2, j_2, k_2) , and (i_3, j_2, k_3) take -1 , and all the other cells are zero.

$j \setminus i$	$k = k_1$	$k = k_2$	$k = k_3$
	$i = i_1$	$i = i_2$	$i = i_3$

$j = j_1$	1		-1
$j = j_2$	-1		1

$j = j_1$	-1	1	
$j = j_2$	1	-1	

$j = j_1$		-1	1
$j = j_2$		1	-1

$$323_6(i_1 i_2 i_3, j_1 j_2, k_1 k_2 k_3) = 222_4(i_1 i_3, j_1 j_2, k_1 k_2) + 222_4(i_2 i_3, j_1 j_2, k_2 k_3)$$

Markov move of degree 6

$332_6(i_1 i_2 i_3, j_1 j_2 j_3, k_1 k_2)$

$i \setminus k$	$j = j_1$	$j = j_2$	$j = j_3$	$j = j_1$	$j = j_2$	$j = j_3$	$j = j_1$	$j = j_2$	$j = j_3$
$k = k_1$	1		-1	-1	1			-1	1
$k = k_2$	-1		1	1	-1			1	-1

Markov move of degree 6

$332_6(i_1 i_2 i_3, j_1 j_2 j_3, k_1 k_2)$

$i \setminus k$	$j = j_1$	$j = j_2$	$j = j_3$	$j = j_1$	$j = j_2$	$j = j_3$	$j = j_1$	$j = j_2$	$j = j_3$
$k = k_1$	1		-1	-1	1			-1	1
$k = k_2$	-1		1	1	-1			1	-1

$i \setminus j$	$j = j_1$	$j = j_2$	$j = j_3$	$j = j_1$	$j = j_2$	$j = j_3$
$k = k_1$	1		-1	-1		1
$k = k_2$	-1	1		1	-1	
$i = i_1$		-1	1		1	-1
$i = i_2$		1		1	-1	
$i = i_3$		-1	1		1	-1

$$332_6(i_1 i_2 i_3, j_1 j_2 j_3, k_1 k_2) = 222_4(i_1 i_2, j_1 j_3, k_1 k_2) + 222_4(i_2 i_3, j_2 j_3, k_1 k_2)$$

Markov move of degree 8, 9, 10

$$\begin{aligned}244_8(i_1 i_2, j_1 j_2 j_3 j_4, k_1 k_2 k_3 k_4) \\ &= 222_4(i_1 i_2, j_1 j_2, k_1 k_2) + 222_4(i_1 i_2, j_3 j_4, k_3 k_4) \\ &\quad + 222_4(i_1 i_2, j_2 j_3, k_3 k_1) \\ 334_8(i_1 i_2 i_3, j_1 j_2 j_3, k_1 k_2 k_3 k_4) \\ &= 222_4(i_1 i_2, j_1 j_2, k_1 k_2) + 222_4(i_1 i_3, j_1 j_2, k_2 k_3) \\ &\quad + 222_4(i_2 i_3, j_2 j_3, k_3 k_4) \\ 344_9(i_1 i_2 i_3, j_1 j_2 j_3 j_4, k_1 k_2 k_3 k_4) \\ &= 222_4(i_1 i_2, j_1 j_4, k_1 k_2) + 222_4(i_1 i_3, j_1 j_2, k_2 k_4) \\ &\quad + 222_4(i_2 i_3, j_1 j_4, k_2 k_4) + 222_4(i_2 i_3, j_3 j_4, k_3 k_4) \\ 344_{10}(i_1 i_2 i_3, j_1 j_2 j_3 j_4, k_1 k_2 k_3 k_4) \\ &= 222_4(i_1 i_3, j_3 j_4, k_1 k_2) + 222_4(i_2 i_3, j_1 j_2, k_3 k_4) \\ &\quad + 222_4(i_1 i_2, j_3 j_4, k_2 k_3) + 222_4(i_1 i_2, j_1 j_4, k_3 k_4)\end{aligned}$$

Sturmfels: Markov basis for $I \times J \times 2$.

Theorem

The set of $I \times J \times 2$ contingency tables has 1-neighbourhood property.

Aoki-Takemura: Markov basis for $3 \times 3 \times K$.

Theorem

The set of $3 \times 3 \times 3$ contingency tables has 2-neighbourhood property and the set of $3 \times 3 \times K$ contingency tables has 3-neighbourhood property for $K \geq 4$.

Theorem

Let N be $3 \times 3 \times 3$ contingency table with $N_{111} = 0$, $N_{211} > 0$, $N_{121} > 0$, and $N_{112} > 0$. N is transmitted to some N' with $N'_{111} = 1$ by at least one of the following Markov moves if and only if there is a contingency table H which has the same marginals as N such that $H_{111} > 0$.

$222_4(12, 12, 12)$, $222_4(12, 12, 13)$, $222_4(12, 13, 12)$,
 $222_4(13, 12, 12)$, $222_4(13, 13, 12)$, $233_6(12, 132, 123)$,
 $323_6(132, 12, 123)$, $332_6(132, 123, 13)$,
 $222_4(13, 13, 32) + 222_4(12, 12, 13)$,
 $222_4(13, 32, 13) + 222_4(12, 13, 12)$,
 $222_4(32, 13, 13) + 222_4(13, 12, 12)$,
 $222_4(23, 23, 23) + 222_4(12, 12, 12)$,
 $222_4(32, 13, 13) + 233_6(13, 132, 123)$

Theorem

The set of $3 \times 4 \times 4$ contingency tables has 3-neighbourhood property.

Theorem

Suppose that $3 \leq I \leq J \leq K$. If the set of $I \times J \times K$ contingency tables has r -neighborhood property then $r \geq I - 1$, and in addition if $I \neq J$ or $J \neq K$ then $r \geq I$.

Let $\mathcal{F}^u(H)$ the subset of $\mathcal{F}(H)$ consisting H with $H_{111} = u$, and $\mathcal{F}^+(H)$ the subset of $\mathcal{F}(H)$ consisting H with $H_{111} > 0$. Similarly let \mathcal{B}^u and \mathcal{B}^+ be the subset consisting M with $M_{111} = u$ and $M_{111} > 0$, respectively, for a Markov basis \mathcal{B} . For the convenience, we assume that the zero table lies in \mathcal{B} .

We write $\mathcal{F}(H)$, $\mathcal{F}^u(H)$ and $\mathcal{F}^+(H)$ by \mathcal{F} , \mathcal{F}^u and \mathcal{F}^+ respectively for short.

For $I \times J \times K$ tables H and H' we denote by $H \geq H'$ if $H_{ijk} \geq H'_{ijk}$ for each i, j, k . An operation $F = M^{(1)} \triangleright M^{(2)} \triangleright \dots \triangleright M^{(s)}$ is said to be applicable for H if all

$$H + M^{(1)}, H + M^{(1)} + M^{(2)}, \dots, H + M^{(1)} + M^{(2)} + \dots + M^{(s)}$$

lie in $\mathcal{F}(H)$. For this operation T , we define $N(T)$ as a table whose (i, j, k) cell has

$$\max_{u=1, \dots, s} \left(- \sum_{a=1}^u M_{ijk}^{(a)}, 0 \right).$$

Note that $N(F)_{111} = 0$ if $M^{(1)}, M^{(2)}, M^{(s-1)} \in \mathcal{F}^0$ and $M^{(s)} \in \mathcal{F}^+$.

The following lemma is one of keys:

Lemma

Let $H \in \mathcal{F}^0$, $M_1^0, \dots, M_r^0 \in \mathcal{B}^0$ and $M_{r+1}^+ \in \mathcal{B}^+$. For an operation $T = M_1^0 \triangleright \dots \triangleright M_r^0 \triangleright M_{r+1}^+$,

T is applicable for $H \iff H \geq N(T)$.

Lemma

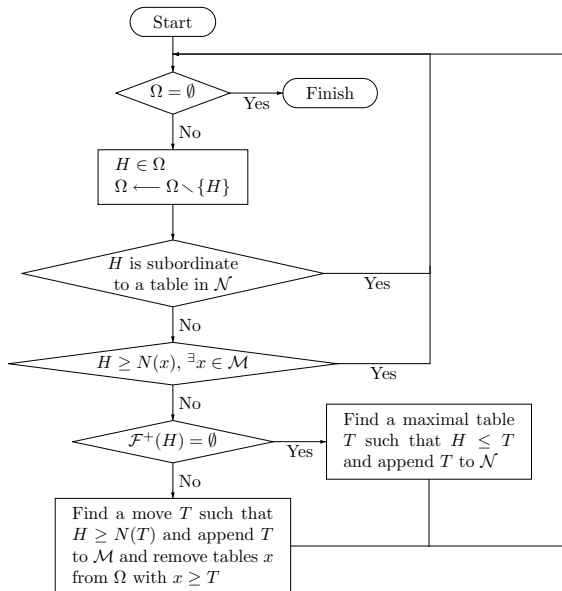
- ① Let H be a table accessible to a table H' with $H'_{111} > 0$. If $G \geq H$ then G is also accessible to a table H'' with $H''_{111} > 0$.
- ② Let H be a table not accessible to any table H' with $H'_{111} > 0$. If $G \leq H$ then G is also not accessible to any table H'' with $H''_{111} > 0$.

By this lemma we only need the set of minimal tables for movability.

Theorem

The following two claims are equivalent.

- ① \mathcal{F} has a r -neighbourhood property.
- ② $\mathcal{F}^+(N)$ intersects with the r -neighbourhood of $N := N(M_1^0 \triangleright \dots \triangleright M_r^0 \triangleright M_{r+1}^+)$ for any M_1^0, \dots, M_r^0 of \mathcal{B}^0 and any M_{r+1}^+ of \mathcal{B}^+ , but $\mathcal{F}^+(\widehat{N})$ does not intersect with the r -neighbourhood of $\widehat{N} := N(\widehat{M}_1^0 \triangleright \dots \triangleright \widehat{M}_{r-1}^0 \triangleright \widehat{M}_r^+)$ for some $\widehat{M}_1^0, \dots, \widehat{M}_{r-1}^0$ of \mathcal{B}^0 and some \widehat{M}_r^+ of \mathcal{B}^+ .



Definition

Let $H = (H_{ijk})$ and $H' = (H'_{ijk})$ be an $I \times J \times K$ table and an $I \times J \times K'$ table, respectively, with $H_{111} = H'_{111} = 0$. We call that H is K -subordinated to H' if there is a partition $\mathbb{P} = P_1, \dots, P_{K'}$ on $\{1, 2, \dots, K\}$ such that

- 1 $P_1 \sqcup \dots \sqcup P_{K'} = \{1, 2, \dots, K\}$,
- 2 $P_k \neq \emptyset$ for any k , and
- 3 $\sum_{\ell \in P_k} H_{ij\ell} \leq H'_{ijk}$ for any i, j, k .

We define ' I -subordinated' and ' J -subordinated' similarly.

Note that $K' \leq K$ and that H is K -subordinated to H itself.

Example

010	100	310	210		010	100	620
101	203	200	100	is K -subordinated to	101	203	400
020	003	121	120		020	004	341

Definition

Let $H = (H_{ijk})$ and $H' = (H'_{ijk})$ be an $I \times J \times K$ table and an $I' \times J' \times K'$ table, respectively, with $H_{111} = H'_{111} = 0$. We call that H is *subordinated* to H' if there are tables G and G' such that H is I -subordinated to G , G is J -subordinated to G' , and G' is K -subordinated to H' .

Since the subordination does not depend on the order of I -, J -, K -subordination we have the following theorem.

Theorem

If H is subordinated to $H' \in \mathcal{F}^+(H') = \emptyset$ then $\mathcal{F}^+(H) = \emptyset$.

This theorem is important for detecting non-movableness by smaller tables.

Theorem (Sumi and Sakata (2009b))

Let H be a $3 \times 3 \times K$ contingency table. If $\mathcal{F}^+(H) = \emptyset$ then H is subordinated to one of the following tables and their permuting tables for permutations preserving 1 on each coordinate:

$$\begin{array}{cc} 0* & ** \\ ** & *0 \end{array}, \quad (2a)$$

$$\begin{array}{ccc} 0*0 & *** & **0 \\ *** & *0* & *00 \\ 0** & 00* & *** \end{array}, \quad (3a)$$

$$\begin{array}{ccc} 0*0 & *** & 0** \\ *** & *0* & 00* \\ **0 & *00 & *** \end{array}, \quad (3b)$$

$$\begin{array}{ccc} 0** & *** & 00* \\ *** & *00 & *0* \\ 0*0 & **0 & *** \end{array}, \quad (3c)$$

$$\begin{array}{ccc} 0*0 & *** & 0** \\ **0 & *00 & *** \\ *** & *0* & 00* \end{array}, \quad (3d)$$

$$\begin{array}{ccc} 0** & *** & 0*0 \\ *** & *00 & **0 \\ 00* & *0* & *** \end{array}. \quad (3e)$$

Here $*$ means a sufficient large integer which is sufficient to be $\max_{i,j,k} H_{ijk}$ for H .

Example

010	100	310	210		010	100	620
101	203	200	100	is K -subordinated to	101	203	400
020	003	121	120		020	004	341

which is a table of type (3a), and thus

$$\mathcal{F}^+ \left(\begin{array}{cccc} 010 & 100 & 310 & 210 \\ 101 & 203 & 200 & 100 \\ 020 & 003 & 121 & 120 \end{array} \right) = \emptyset.$$

Non-movable tables $\mathfrak{N}(1, 1, 1)$

Let $\mathfrak{N}(1, 1, 1)$ be the set consisting of tables of type (2a), (3a)–(3e) and the following 26 tables and their permuting tables for permutations preserving 1 on each coordinate:

0**0 *** **0	*** *00* *000	0*** 000* *0*	00*0 00** ***	0**0 *** 0**0	*** *00* 000*	***0 *000 **0*	00*0 *0*0 ***	0*** ***0 0*00	*** *000 **00	000* *** 0*0*	00** *0*0 ***
0**0 *** 0*00	**0* *000 ***	*** *00* 0*0*	00*0 *** 0***	0**0 *** 0*00	*** *00* **00	00** *000 ***	00*0 *0** ***0	0*00 *** **00	*** *0*0 *000	**0* *0** 0000	0*0* 00** ***
0**0 *** **0*	**0* *000 *00*	0**0 00*0 ***	0*0* *** 000*	0*00 *** 0**0	*** 000* 00**	***0 *0** 00*0	**00 *00* ***	0*00 ***0 0***	**00 *000 ***	*** *0*0 00**	0*0* *** 000*
0**10 *** 0**0	*** *000 *0*0	0*** *00* 00*0	0*0* *00* ***	0*10 *** 0**0	*** *000 *00*	0*** *0** 00**	0*0* *0*0 ***	0*00 *** 1**0	*** *000 *0*0	0*** *0** 00*0	0*0* *00* ***
0**00 *** 0**0	*** *000 *0*0	1*0* *00* ***	0*** *0** 00*0	0**0 *** 0**1	*** *000 *0*0	0*** *0** 00*0	0*0* *00* ***	0*00 *** 0*0*	*** *000 *0*0	1*** *0** 00*0	0*0* *00* ***
0**00 *** 0**0	*** *000 *0*0	0*** *0** 00*0	0*0* *01* ***	0**0 *** 0**0	*** *000 *0*0	0*** *0** 00*0	0*0* *01* ***	0*00 *** 0*0*	*** *000 *0*0	**0** *01** 00*0	0*0* *00* ***
0**00 *** 0**0	*** *000 *0*0	0*** *0** 10*0	0*0* *00* ***	0**0 *** 0**0	*** *000 *0*0	0*** *0** 00*1	0*0* *00* ***	0*00 *** 0*0*	*** *001 *0*0	0*** *0** 00*0	0*0* *00* ***
0**00 *** 0**0	*** *000 *1*0	0*** *0** 00*0	0*0* *00* ***	0**0 *** 0**0	*** *000 *0*1	0*** *0** 00*0	0*0* *00* ***	0*00 *** 0**0	*** *000 *0*0	0*** *0** 01*0	0*0* *00* ***
0**00 *** 0**0	*** *000 *0*0	0*** *0** 00*0	0*1* *00* ***	0**0 *** 0**0	*** *010 *0*0	0*** *0** 00*0	0*0* *00* ***				

Theorem

A contingency table H is not reachable to a table H' with $H'_{111} > 0$ by Markov moves if and only if H is subordinate of a table in $\mathfrak{N}(1, 1, 1)$.

Movable tables $\mathfrak{M}(1, 1, 1) \mid$

By $M = M^{(1)} \triangleright M^{(2)} \triangleright \dots \triangleright M^{(s)}$ we denote an operation by transformations plussing $M^{(1)}$, next $M^{(2)}$, step by step, and finally $M^{(s)}$. Hereafter, the symbol $pqr_d(\dots)$ means a move with degree d of the minimal Markov basis obtained by Aoki and Takemura (2003), which is essentially a $p \times q \times r$ table.

Let $\mathfrak{M}(1, 1, 1)$ be the set consisting of the below 105 tables and their permuting tables for permutations preserving 1 on each coordinate:

» skip tables

$222_4(12, 12, 12)$, $233_6(12, 132, 123)$, $332_6(123, 213, 21)$,
 $323_6(132, 12, 123)$, $343_8(312, 4321, 321)$, $343_8(312, 2314, 132)$,
 $343_8(213, 1342, 213)$, $343_8(123, 1234, 123)$, $334_8(312, 321, 3421)$,
 $334_8(123, 123, 1243)$, $334_8(312, 132, 2413)$, $334_8(213, 213, 1432)$,
 $344_9(132, 3241, 2134)$, $344_9(123, 1432, 1342)$,
 $344_{10}(231, 3421, 1234)$, $344_{10}(132, 3412, 1324)$,
 $344_9(132, 2134, 3241)$, $344_{10}(123, 2341, 2341)$,

Movable tables $\mathfrak{M}(1, 1, 1)$ II

$344_{10}(132, 1324, 3412)$, $344_{10}(231, 1243, 4321)$,
 $244_8(12, 1342, 1234)$,
 $222_4(13, 13, 32) \triangleright 222_4(12, 12, 13)$, $222_4(32, 23, 32) \triangleright 222_4(12, 12, 12)$,
 $222_4(13, 32, 13) \triangleright 222_4(12, 13, 12)$, $222_4(23, 13, 31) \triangleright 222_4(13, 12, 12)$,
 $222_4(13, 32, 12) \triangleright 233_6(12, 143, 123)$,
 $222_4(13, 34, 43) \triangleright 233_6(12, 132, 124)$,
 $222_4(32, 24, 34) \triangleright 233_6(12, 132, 124)$,
 $222_4(13, 13, 42) \triangleright 233_6(12, 142, 143)$,
 $222_4(32, 34, 24) \triangleright 233_6(12, 142, 123)$,
 $222_4(13, 23, 31) \triangleright 233_6(12, 143, 124)$,
 $222_4(23, 13, 41) \triangleright 233_6(13, 142, 123)$,
 $222_4(13, 13, 32) \triangleright 332_6(123, 213, 31)$,
 $222_4(13, 14, 32) \triangleright 332_6(132, 123, 13)$,
 $222_4(12, 32, 13) \triangleright 332_6(123, 314, 21)$,
 $222_4(12, 13, 42) \triangleright 323_6(132, 12, 143)$,
 $222_4(12, 23, 34) \triangleright 323_6(123, 21, 214)$,
 $222_4(13, 32, 13) \triangleright 323_6(123, 31, 214)$,
 $222_4(13, 13, 42) \triangleright 343_8(312, 4321, 341)$,

Movable tables $\mathfrak{M}(1, 1, 1)$ III

$222_4(12, 13, 42) \triangleright 343_8(123, 1234, 134),$
 $222_4(13, 12, 32) \triangleright 343_8(312, 2314, 143),$
 $222_4(13, 13, 42) \triangleright 343_8(213, 1342, 413),$
 $222_4(13, 13, 42) \triangleright 343_8(132, 2143, 413),$
 $222_4(13, 13, 32) \triangleright 343_8(312, 2314, 143),$
 $222_4(13, 14, 42) \triangleright 343_8(132, 2143, 413),$
 $222_4(12, 43, 23) \triangleright 343_8(213, 1342, 214),$
 $222_4(32, 14, 42) \triangleright 343_8(213, 1342, 213),$
 $222_4(23, 12, 31) \triangleright 343_8(213, 2314, 142),$
 $222_4(12, 32, 13) \triangleright 343_8(123, 1324, 124),$
 $222_4(12, 14, 42) \triangleright 334_8(132, 213, 4132),$
 $222_4(12, 24, 43) \triangleright 334_8(213, 213, 1342),$
 $222_4(23, 34, 42) \triangleright 334_8(312, 421, 4321),$
 $222_4(13, 23, 41) \triangleright 334_8(312, 431, 3421),$
 $222_4(13, 23, 41) \triangleright 334_8(123, 134, 1234),$
 $222_4(13, 23, 41) \triangleright 334_8(312, 143, 2413),$
 $222_4(13, 32, 14) \triangleright 334_8(213, 314, 1432),$
 $222_4(13, 23, 21) \triangleright 334_8(312, 143, 2413),$

Movable tables $\mathfrak{M}(1, 1, 1)$ IV

$222_4(12, 34, 41) \triangleright 334_8(123, 124, 1243),$
 $222_4(13, 23, 41) \triangleright 344_{10}(231, 4231, 1243),$
 $222_4(13, 14, 42) \triangleright 344_{10}(231, 1243, 3241),$
 $222_4(13, 43, 34) \triangleright 344_{10}(123, 2341, 2431),$
 $233_6(13, 124, 321) \triangleright 222_4(12, 13, 13),$
 $233_6(13, 123, 431) \triangleright 222_4(12, 13, 12),$
 $233_6(13, 134, 423) \triangleright 222_4(12, 12, 14),$
 $233_6(32, 423, 432) \triangleright 222_4(12, 12, 12),$
 $233_6(13, 234, 413) \triangleright 222_4(12, 13, 12),$
 $233_6(32, 143, 143) \triangleright 222_4(13, 12, 12),$
 $233_6(13, 123, 421) \triangleright 233_6(12, 143, 143),$
 $233_6(13, 123, 321) \triangleright 233_6(12, 143, 124),$
 $233_6(13, 134, 423) \triangleright 332_6(123, 213, 41),$
 $233_6(13, 134, 423) \triangleright 343_8(132, 2143, 413),$
 $233_6(13, 123, 421) \triangleright 233_6(12, 142, 143),$
 $343_8(213, 4123, 431) \triangleright 222_4(12, 13, 12),$
 $343_8(213, 4123, 341) \triangleright 323_6(132, 13, 124),$
 $343_8(213, 4123, 421) \triangleright 233_6(12, 143, 123),$

Movable tables $\mathfrak{M}(1, 1, 1) \vee$

$334_8(213, 134, 2143) \triangleright 222_4(12, 12, 14),$
 $334_8(213, 341, 3124) \triangleright 222_4(12, 12, 14),$
 $334_8(213, 321, 4123) \triangleright 233_6(12, 142, 134),$
 $334_8(213, 341, 3124) \triangleright 332_6(123, 214, 41),$
 $344_9(132, 2413, 2314) \triangleright 222_4(13, 12, 12),$
 $344_9(123, 4213, 3124) \triangleright 222_4(12, 13, 12),$
 $344_9(123, 4123, 3214) \triangleright 222_4(12, 12, 14),$
 $222_4(13, 23, 31) \triangleright 222_4(23, 34, 23) \triangleright 222_4(12, 13, 12),$
 $222_4(12, 14, 32) \triangleright 222_4(23, 13, 21) \triangleright 222_4(13, 12, 13),$
 $222_4(13, 13, 42) \triangleright 222_4(32, 23, 34) \triangleright 222_4(12, 12, 14),$
 $222_4(12, 32, 13) \triangleright 222_4(32, 12, 14) \triangleright 222_4(13, 13, 12),$
 $222_4(12, 13, 32) \triangleright 222_4(13, 14, 43) \triangleright 222_4(12, 12, 14),$
 $222_4(12, 34, 43) \triangleright 222_4(13, 13, 42) \triangleright 222_4(12, 12, 14),$
 $222_4(13, 13, 42) \triangleright 222_4(32, 42, 43) \triangleright 222_4(12, 12, 14),$
 $222_4(12, 24, 34) \triangleright 222_4(32, 23, 42) \triangleright 222_4(12, 12, 12),$
 $222_4(12, 43, 23) \triangleright 222_4(32, 23, 42) \triangleright 222_4(12, 12, 12),$
 $222_4(13, 14, 43) \triangleright 222_4(23, 13, 31) \triangleright 222_4(13, 12, 12),$
 $222_4(12, 43, 34) \triangleright 222_4(13, 32, 14) \triangleright 222_4(12, 13, 12),$
 $222_4(12, 14, 42) \triangleright 222_4(32, 13, 13) \triangleright 222_4(13, 12, 14),$

$222_4(13, 14, 42) \triangleright 222_4(13, 32, 13) \triangleright 222_4(12, 13, 14),$
 $222_4(13, 32, 13) \triangleright 222_4(23, 34, 24) \triangleright 222_4(12, 13, 12),$
 $222_4(13, 34, 31) \triangleright 222_4(32, 13, 14) \triangleright 222_4(13, 12, 12),$
 $222_4(12, 32, 13) \triangleright 222_4(13, 34, 41) \triangleright 222_4(12, 14, 12),$
 $222_4(12, 23, 31) \triangleright 222_4(23, 14, 41) \triangleright 222_4(13, 13, 12),$
 $222_4(12, 34, 43) \triangleright 222_4(13, 13, 42) \triangleright 332_6(123, 213, 41),$
 $222_4(13, 13, 42) \triangleright 222_4(13, 23, 41) \triangleright 233_6(12, 143, 143),$
 $222_4(12, 14, 32) \triangleright 222_4(13, 13, 43) \triangleright 332_6(123, 213, 41),$
 $222_4(13, 14, 42) \triangleright 222_4(12, 23, 31) \triangleright 332_6(132, 134, 14),$
 $222_4(12, 14, 42) \triangleright 233_6(13, 123, 341) \triangleright 222_4(12, 13, 12),$
 $222_4(12, 32, 13) \triangleright 233_6(13, 134, 421) \triangleright 222_4(12, 12, 14).$

The image $N(\mathfrak{M}(1, 1, 1))$ by N gives a necessary condition for a table H with $H_{111} = 0$ to reach to a table H' with $H'_{111} > 0$ as follows.

Theorem

Suppose that \mathcal{F}_t is obtained from the previous frame \mathcal{F}_{t-1} by adding 1 at the $(1, 1, 1)$ -cell. Let ϕ be a map from \mathcal{F}_{t-1} to \mathcal{F}_t by simply adding 1 at the $(1, 1, 1)$ -cell. Then

$$\begin{aligned} \mathcal{F}_t = \{ & \phi(H) \mid H \in \mathcal{F}_{t-1} \} \\ & \cup \{ \phi(H) - F \mid H \in \mathcal{F}_{t-1}, F \in \mathfrak{M}(1, 1, 1) \}. \end{aligned}$$

For a give marginals α, β, γ , we have an algorithm to obtain $\mathcal{F}(\alpha, \beta, \gamma)$.
In particular we decide whether $\mathcal{F}(\alpha, \beta, \gamma)$ is empty or not.

For a $3 \times 4 \times 4$ contingency table H , we determine the types of H such that the set $\mathcal{F}(H)$ of contingency tables with the marginals as same as H has no table T with $T_{111} > 0$, i.e. $\mathcal{F}^+(H) = \emptyset$. For given i, j, k , we give a (minimal) Markov basis \mathcal{B} for the set \mathcal{F} of $3 \times 4 \times 4$ contingency table with the property that if \mathcal{F} has a table T' with $T'_{ijk} = 0$ and a table T with $T_{ijk} > 0$, then for any table T with $T_{ijk} > 0$ of \mathcal{F} , there is a move $M \in \mathcal{B}$ such that $T + M \in \mathcal{F}$ with $(T + M)_{ijk} < T_{ijk}$.

▶ Go to thanks

- Agresti, A., An introduction to Categorical Data Analysis. *John Wiley & Sons, Inc.*, 1996.
- Aoki, S., Exact methods and Markov chain Monte Carlo methods of conditional inference for contingency tables, Doctor Thesis, Tokyo University 2004.
- Aoki, S. and Takemura, A., *Minimal basis for connected Markov chain over $3 \times 3 \times K$ contingency tables with fixed two-dimensional marginals*, Australian and New Zealand Journal of Statistics, 45, 229–249, 2003.
- Saito, M., Sturmfels, B. and Takayama, N., Hypergeometric polynomials and integer programming. *Compositio Mathematica* 115, 185–204, 1997.

- Sturmfels, B., Gröbner bases and convex polytopes. *American Mathematical Society, University Lecture Series* 8, 1996.
- Sakata, T. and Sumi, T. (2008): Lifting between the sets of three-way contingency tables and r -neighborhood property. *Electronic Proceedings of COMPSTAT '2008, Contributed Papers, Categorical Data Analysis*, 87–94.
- Sumi, T. and Sakata, T. (2009a): A proof of 2-neighborhood theorem for $3 \times 3 \times 3$ tables. *preprint*.
- Sumi, T. and Sakata, T. (2009b): The set of $3 \times 3 \times K$ contingency tables for $K \geq 4$ has 3-neighborhood property. *preprint*.

Thank you for your attention.