数理解析研究所講究録 1325 巻 2003 年 215-220

$m \times n$ 分割表の近似数え上げスキームの提案

東京大学大学院 情報理工学系研究科 数理情報学専攻
来嶋 秀治 (Shuji Kijima)
松井 知己 (Tomomi Matsui)

Department of Mathematical Informatics,
Graduate School of Information Science and Technology,
University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan

概 要

本報告では $m \times n$ 分割表に対する新しい近似数え上げ手法を提案する。この手法は、Dyer and Greenhill が 2 行分割表に対して提案した手法を改良し、拡張したものである。我々は近似解の精度を保証するだけでなく、得られた推定量の厳密解からの偏りの大きさについても議論する。

1 はじめに

2 元分割表は正の整数からなる行和と列和を持ち、表中の値として非負整数をとる表（行列）である。2 元分割表は医療統計の分野などで統計データを扱うのに用いられる。与えられた行和および列和を満たす 2 元分割表の個数を厳密に数える問題は、行数が 2 の時でさえ#P 完全であることが知られている [4]。

2000 年に Dyer and Greenhill によって MCMC（マルコフ連鎖モンテカルロ）法を用いた $2 \times n$ 分割表の個数を求める近似解法が提案された [3]。この方法は非常に直感的だが、$m \times n$ 分割表に適用した場合、精度や偏りの理論的考察が困難となる。

実際、我々は Dyer and Greenhill の手法と我々の手法を実装し、2 行分割表に対して厳密解とそれぞれの解法で得られた近似解の平均値との比較を行った。図 1 は行和 $(42, 38)$、列和 $(10, \ldots, 10)$ の周辺和をもつ 2×8 分割表に対して、それぞれの解法を 2,000 回ずつ行って得られた近似解のヒストグラムである。この問題の厳密解は $9,162,736$ であるが、Dyer and Greenhill の手法による近似解の平均は大きく偏ったのに対して、我々の手法による近似解にはほとんど偏りは認めるられなかった。

本報告では、分割表の個数に関する性質を考慮することで Dyer and Greenhill の方法を改良する。さらに我々の方法は $m \times n$ 分割表に拡張することが可能である。本報告では、我々の手法によって得られる推定量の精度とその期待値の厳密解からの偏りの大きさについて議論する。

2 近似解法

整数（非負整数、正整数）全体の集合をそれぞれ $\mathbb{Z}$ ($\mathbb{Z}_+$、$\mathbb{Z}_{++}$) で表すことにする。ベクトル $\boldsymbol{r} = (r_1, \ldots, r_m) \in \mathbb{Z}_+^m$ と $\boldsymbol{s} = (s_1, \ldots, s_n) \in \mathbb{Z}_+^m$ は正整数 $N \in \mathbb{Z}_{++}$ に対して、$\sum_{i=1}^m r_i = \sum_{j=1}^n s_j = N$ を満たすとする。行和 $\boldsymbol{r}$ および列和 $\boldsymbol{s}$ をもち、非負整数を表値にとる m 行 n 列の 2 元分割表全体

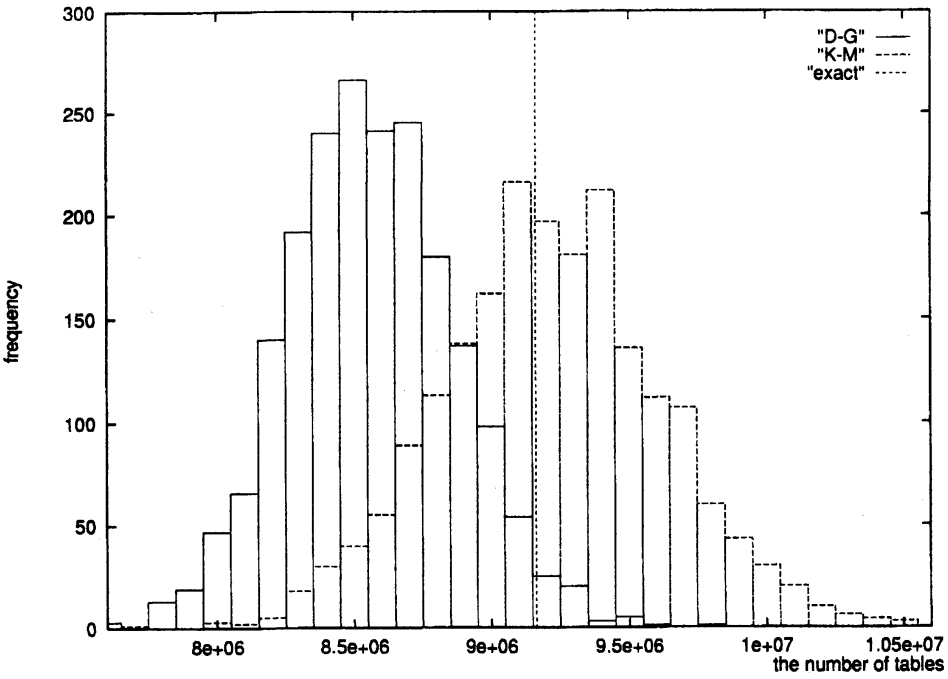

図 1: The histogram

の集合 $\Sigma_{\boldsymbol{r},\boldsymbol{s}}$ を、

$$\Sigma_{\boldsymbol{r},\boldsymbol{s}} \stackrel{\text{def.}}{=} \left\{ X \in \mathbb{Z}_+^{m\times n} \;\middle|\; \sum_{j=1}^{n} X_{ij} = r_i \quad (1 \le \forall i \le m), \quad \sum_{i=1}^{m} X_{ij} = s_j \quad (1 \le \forall j \le n) \right\}$$

で定義する。但し、X_{ij} は分割表 X の i 行 j 列の値を表す。明らかに分割表の個数 $|\Sigma_{\boldsymbol{r},\boldsymbol{s}}|$ は行と列を置き換えても変わらない。したがって、一般性を失うことなく $m \le n$ とする。部分集合 $\Omega_i \subset \Sigma_{\boldsymbol{r},\boldsymbol{s}}$ を $\Omega_i = \{X \in \Sigma_{\boldsymbol{r},\boldsymbol{s}} | X_{in} \ge \lceil s_n/m \rceil\}$ で定義する。また、添え字 k は $r_k = \max\{r_1, \ldots, r_m\}$ を満たすとする。この時、

$$\begin{aligned} \tilde{\boldsymbol{r}} &\stackrel{\text{def.}}{=} (r_1, \ldots, r_k - \lceil s_n/m \rceil, \ldots, r_m), \\ \tilde{\boldsymbol{s}} &\stackrel{\text{def.}}{=} \begin{cases} (s_1, \ldots, s_{n-1}, \lfloor \frac{m-1}{m} s_n \rfloor), & s_n > 1, \\ (s_1, \ldots, s_{n-1}), & s_n = 1, \end{cases} \end{aligned}$$

を定義する。明らかに $|\Omega_k| = |\Sigma_{\tilde{\boldsymbol{r}},\tilde{\boldsymbol{s}}}|$ が成り立つ。もし、$\rho = |\Omega_k|/|\Sigma_{\boldsymbol{r},\boldsymbol{s}}|$ と $|\Sigma_{\tilde{\boldsymbol{r}},\tilde{\boldsymbol{s}}}|$ がわかれば $|\Sigma_{\boldsymbol{r},\boldsymbol{s}}| = |\Sigma_{\tilde{\boldsymbol{r}},\tilde{\boldsymbol{s}}}|/\rho$ を計算できる。しかし、一般に ρ および $|\Sigma_{\tilde{\boldsymbol{r}},\tilde{\boldsymbol{s}}}|$ を求めることは困難である。いま、$\Sigma_{\boldsymbol{r},\boldsymbol{s}}$ 上で一様標本抽出が可能ならば、ρ の値の推定にモンテカルロ法を適用できる。すなわち M 回の標本抽出を行い、U 個の分割表が Ω_k に含まれるならば、$(U+1)/(M+1)$ を ρ の推定量とする。こうして、もとの問題に比べてサイズの小さな分割表 $\Sigma_{\tilde{\boldsymbol{r}},\tilde{\boldsymbol{s}}}$ の数え上げ問題に帰着させることが出来る。この手続きを繰り返して、分割表のサイズが 2 行 2 列になるまで小さくする。2×2 分割表の個数は定数時間で求めることができるので、再帰的に $|\Sigma_{\boldsymbol{r},\boldsymbol{s}}|$ を求めることができる。

3 分割表の個数に関する諸定理

我々のスキームでは、問題のサイズを小さくする目的で部分集合 Ω_k を定義した。この定義には二つの重要な意味がある。一つは行の添え字 k の決め方で、もう一つはサイズ縮小の値 $\lceil s_n/m \rceil$ である。これは二つの背反な要求から来る。まず、多項式時間のアルゴリズムを得るために、効率的に問題のサイズ縮小を行う必要がある。また、比の値 ρ が小さすぎると ρ の推定量が誤差に敏感になるので ρ はある程度十分大きくなければならない。この目的で我々は次の定理を得た。

定理 1 添え字 $k \in \{1,\ldots,m\}$ は $r_k = \max\{r_1,\ldots,r_m\}$ を満たすとする。この時、$|\Omega_k| \geq \frac{1}{m}|\Sigma_{\boldsymbol{r},\boldsymbol{s}}|$ が成り立つ。

定理 1 から、我々のスキームに現れる理論比 ρ は $\rho \geq 1/m$ を満たすことが分かる。以下、定理 1 を示す手順について説明する。

まず、分割表の個数について以下の 2 つの補題を示すことができる。

補題 1 ベクトル $\boldsymbol{r},\boldsymbol{r}' \in \mathbb{Z}_+^2$ および $\boldsymbol{s} \in \mathbb{Z}_+^n$ は $|r_1 - r_2| \leq |r_1' - r_2'|, \sum_{i=1}^2 r_i = \sum_{i=1}^2 r_i' = \sum_{j=1}^n s_j = N$ を満たすとする。この時、集合 $|\Sigma_{\boldsymbol{r},\boldsymbol{s}}|$ と $|\Sigma_{\boldsymbol{r}',\boldsymbol{s}}|$ について $|\Sigma_{\boldsymbol{r},\boldsymbol{s}}| \geq |\Sigma_{\boldsymbol{r}',\boldsymbol{s}}|$ が成り立つ。

補題 2 ベクトル $\boldsymbol{r},\boldsymbol{r}' \in \mathbb{Z}_+^m$ および $\boldsymbol{s} \in \mathbb{Z}_+^n$ は $|r_1 - r_2| \leq |r_1' - r_2'|$, $r_i = r_i'\ (i = 3,\ldots,m)$, $\sum_{i=1}^m r_i = \sum_{i=1}^m r_i' = \sum_{j=1}^n s_j = N$ を満たすとする。この時、集合 $|\Sigma_{\boldsymbol{r},\boldsymbol{s}}|$ と $|\Sigma_{\boldsymbol{r}',\boldsymbol{s}}|$ について $|\Sigma_{\boldsymbol{r},\boldsymbol{s}}| \geq |\Sigma_{\boldsymbol{r}',\boldsymbol{s}}|$ が成り立つ。

いま、点 $\boldsymbol{y} = (y_1,\ldots,y_m)$ は $\mathbb{R}^m$ 上の点とする。$N^- = \sum_{j=1}^{n-1} s_j = N - s_n$ として、$\mathbb{R}^m$ 空間上に超平面

$$S \stackrel{\text{def.}}{=} \left\{ \boldsymbol{y} \in \mathbb{R}^m \;\middle|\; \sum_{i=1}^m y_i = N^- \right\}$$

を定義する。また、$\boldsymbol{r} = (r_1,\ldots,r_m)$ に対し、$(m-1)$ 次元単体 S^+ と S_0 をそれぞれ、

$$S^+ \stackrel{\text{def.}}{=} \{\boldsymbol{y} \in S \mid y_i \geq 0,\ i = 1,\ldots,m\}, \tag{1}$$

$$S_0 \stackrel{\text{def.}}{=} \{\boldsymbol{y} \in S \mid r_i - s_n \leq y_i \leq r_i,\ i = 1,\ldots,m\}, \tag{2}$$

で定める。

いま、$\boldsymbol{x} \in \mathbb{R}^m$ を $\boldsymbol{x} \stackrel{\text{def.}}{=} \boldsymbol{r} - \boldsymbol{y}$ と定める。任意の $\boldsymbol{x}$ に対して、$x_i = X_{in} \quad (i = 1,\ldots,m)$ となる分割表 $X \in \Sigma_{\boldsymbol{r},\boldsymbol{s}}$ が存在するための必要十分条件は、

$$0 \leq x_i \leq \min\{r_i, s_n\} \qquad (i = 1,\ldots,m),$$
$$\sum_{i=1}^m x_i = s_n, \qquad \boldsymbol{x} \in \mathbb{Z}^m,$$

である。この条件は $\boldsymbol{x} = \boldsymbol{r} - \boldsymbol{y}$ より、領域 S^+, S_0 を用いて $\exists \boldsymbol{y} \in S^+ \cap S_0 \cap \mathbb{Z}^m$ と表される。

任意の $k,l \in \{1,\ldots,m\},\ k \neq l$ に対して、$\mathbb{R}^m$ 空間上の超平面 H_{kl} を

$$H_{kl} = \{\boldsymbol{y} \in \mathbb{R}^m \mid -y_k + y_l = -r_k + r_l\} \tag{3}$$

で定義する。また、$i \in \{1,\ldots,m\}$ に対して領域 S_i を

$$S_i \stackrel{\text{def.}}{=} \{\boldsymbol{y} \in S_0 \mid r_i - y_i = \max\{r_l - y_l \mid l = 1,\ldots,m\}\} \tag{4}$$

と定義する。さらに、任意の点 $\boldsymbol{y} \in S$ に対して関数 $q(\boldsymbol{y})$ を

$$q(\boldsymbol{y}) = \begin{cases} |\Sigma_{\boldsymbol{y},\boldsymbol{s}^-}| & (\boldsymbol{y} \in \mathbb{Z}_+^m \cap S_0 \cap S^+), \\ 0 & (\text{otherwise}), \end{cases} \tag{5}$$

と定義する。但し、$\boldsymbol{s}^- = (s_1, \ldots, s_{n-1})$ である。明らかに、$|\Sigma_{\boldsymbol{r},\boldsymbol{s}}| = \sum_{\boldsymbol{y} \in S_0} q(\boldsymbol{y})$ が成り立つ。この時、$q(\boldsymbol{y})$ について以下の補題が成り立つ。

補題 3 行の添え字 $k \in \{1, \ldots, m\}$ は $r_k = \max\{r_1, \ldots, r_m\}$ を満たすとする。また、$l \in \{1, \ldots, m\}$, $l \neq k$ に対し $\boldsymbol{y} \in S_k$ の H_{kl} に関して対称な点 $\boldsymbol{y}^* = h_{kl}(\boldsymbol{y})$ は $q(\boldsymbol{y}) \geq q(\boldsymbol{y}^*)$ を満たす。

任意の $i \in \{1, \ldots, m\}$ に対し、関数 $Q_i \stackrel{\text{def.}}{=} \sum_{\boldsymbol{y} \in S_i} q(\boldsymbol{y})$ を定義する。明らかに、$|\Sigma_{\boldsymbol{r},\boldsymbol{s}}| \leq \sum_{i=1}^m Q_i$ が成り立つ。いま、次の補題を得る。

補題 4 行の添え字 k は $r_k = \max\{r_1, \ldots, r_m\}$ を満たすとする。この時、任意の $i \in \{1, \ldots, m\}$ に対して、$Q_k \geq Q_i$ が成り立つ。

補題 4 から、行の添え字 k が $r_k = \max\{r_1, \ldots, r_m\}$ を満たすとき、$|\Sigma_{\boldsymbol{r},\boldsymbol{s}}| \leq \sum_{i=1}^m Q_i \leq mQ_k$ が成り立つ。また、Q_k の定義より

$$\begin{aligned} Q_k &= |\{X \in \Sigma_{\boldsymbol{r},\boldsymbol{s}} \mid \forall i \; X_{kn} \geq X_{in}\}| \\ &\leq |\{X \in \Sigma_{\boldsymbol{r},\boldsymbol{s}} \mid X_{kn} \geq \tfrac{1}{m} s_n\}| \quad = |\Omega_k| \end{aligned}$$

が成り立つ。すなわち $|\Sigma_{\boldsymbol{r},\boldsymbol{s}}| \leq m|\Omega_k|$ となり、定理 1 を得る。

4 推定量の誤差と偏り

ここでは、我々の近似解法で得られる近似解の誤差と偏りについて議論する。提案した近似解法中で $m \times n$ 分割表の一様標本抽出を仮定した。しかし、一般に $m \times n$ 分割表の一様標本抽出は困難であるため、我々の提案するスキームでは近似一様標本抽出を仮定する。集合 Ω 上の分布 π と ν のあいだの総分布距離 $\mathrm{d_{TV}}(\pi, \nu)$ を、$\mathrm{d_{TV}}(\pi, \nu) = \frac{1}{2} \sum_{x \in \Omega} |\pi(x) - \nu(x)|$ で定義する。いま、集合 $\Sigma_{\boldsymbol{r},\boldsymbol{s}}$ 上の一様分布を π で表す。任意の正数 $\varepsilon < 1$ に対して、ある近似的な一様標本抽出法が存在して、標本は $\Sigma_{\boldsymbol{r},\boldsymbol{s}}$ 上の分布 ν に従って抽出されるものとし、分布 ν は総分布距離 $\mathrm{d_{TV}}(\pi, \nu) \leq \varepsilon/(6mR)$ を満たすとする。但し、R は Z を得るために必要な問題のサイズ縮小の回数とする。この標本抽出法を用いて標本数 $M = 108mR^2\varepsilon^{-2} \ln(2R/\delta)$ のモンテカルロ法を行い、前節で提案した手法における比 ρ の推定量 $\hat{\rho}$ を計算する。逐次この推定比を用いて $|\Sigma_{\boldsymbol{r},\boldsymbol{s}}|$ の推定量 Z を求める。すなわち我々のスキームは多項式回の問題縮小と多項式個の標本抽出で終了する。この時、推定量 Z に関して次の二つの定理が成り立つ。

定理 2 推定量 Z は

$$\Pr[(1-\varepsilon)|\Sigma_{\boldsymbol{r},\boldsymbol{s}}| \leq Z \leq (1+\varepsilon)|\Sigma_{\boldsymbol{r},\boldsymbol{s}}|] \geq 1 - \delta$$

を満たす。

定理 2 は [4, 3] と同様な以下の手順で示すことができる。

1. $1 \leq \forall i \leq R$, $\hat{\rho}_i \stackrel{\text{def.}}{=} \mathrm{E}[U_i/M] \geq 1/m - \varepsilon/6mR$,

2. $|\rho_i - \hat{\rho}_i| \leq \dfrac{\varepsilon}{6R-\varepsilon}\hat{\rho}_i$,
3. $\Pr\left[|Z_i - \hat{\rho}_i| > \dfrac{\varepsilon}{6R-\varepsilon}\hat{\rho}_i\right] \leq \dfrac{\delta}{R}$,
4. $1-\delta$ 以上の確率で、$|(Z_1 \cdots Z_R)^{-1} - (\rho_1 \cdots \rho_R)^{-1}| \leq \varepsilon(\rho_1 \cdots \rho_R)^{-1}$.

定理 3 推定量 Z は

$$\frac{|\mathrm{E}[Z] - \Sigma_{\boldsymbol{r},\boldsymbol{s}}|}{|\Sigma_{\boldsymbol{r},\boldsymbol{s}}|} \leq \frac{\varepsilon}{4} + \mathrm{e}^{-90R^3\varepsilon^{-2}\ln(2R/\delta)} < (\frac{1}{4} + \frac{1}{10^{27}})\,\varepsilon$$

を満たす。

証明: $Z_1, \ldots, Z_R$ は独立なので

$$\mathrm{E}[Z] = \sigma\mathrm{E}\left[\prod_{i=1}^{R}\frac{1}{Z_i}\right] = \sigma\prod_{i=1}^{R}\mathrm{E}\left[\frac{1}{Z_i}\right],$$

が成り立つ。いま、

$$\begin{aligned}\mathrm{E}\left[\frac{1}{Z_i}\right] &= \sum_{U_i=0}^{M}\frac{M+1}{U_i+1}\binom{M}{U_i}\hat{\rho}_i^{U_i}(1-\hat{\rho}_i)^{M-U_i} \\ &= \frac{1}{\hat{\rho}_i}\sum_{U_i=0}^{M}\binom{M+1}{U_i+1}\hat{\rho}_i^{U_i+1}(1-\hat{\rho}_i)^{M-U_i} = \frac{1}{\hat{\rho}_i}\left\{1-(1-\hat{\rho}_i)^{M+1}\right\},\end{aligned}$$

である。したがって等式

$$\mathrm{E}[Z] = \sigma\prod_i\frac{1}{\hat{\rho}_i}\left\{1-(1-\hat{\rho}_i)^{M+1}\right\} \tag{6}$$

が成り立つ。式 (6) より、$||\Sigma_{\boldsymbol{r},\boldsymbol{s}}| - \mathrm{E}[Z]|$ は

$$\begin{aligned}||\Sigma_{\boldsymbol{r},\boldsymbol{s}}| - \mathrm{E}[Z]| &= \left|\sigma\prod_i\frac{1}{\rho_i} - \sigma\prod_i\frac{1}{\hat{\rho}_i}\left\{1-(1-\hat{\rho}_i)^{M+1}\right\}\right| \\ &\leq \left|\sigma\prod_i\frac{1}{\rho_i} - \sigma\prod_i\frac{1}{\hat{\rho}_i}\right|\left|\left\{1-(1-\hat{\rho}_i)^{M+1}\right\}\right| + \left|\sigma\prod_i\frac{1}{\rho_i}(1-\hat{\rho}_i)^{M+1}\right|,\end{aligned}$$

を満たす。従って、

$$\begin{aligned}\left|\sigma\prod_i\frac{1}{\rho_i} - \sigma\prod_i\frac{1}{\hat{\rho}_i}\right| &= \sigma\prod_i\frac{1}{\rho_i}\left|1-\prod_i\frac{\rho_i}{\hat{\rho}_i}\right| \leq |\Sigma_{\boldsymbol{r},\boldsymbol{s}}|\left\{\left(1+\frac{\varepsilon}{6R-\varepsilon}\right)^R - 1\right\} \\ &\leq |\Sigma_{\boldsymbol{r},\boldsymbol{s}}|\left\{\exp\left(\frac{\varepsilon R}{6R-\varepsilon}\right) - 1\right\} \leq |\Sigma_{\boldsymbol{r},\boldsymbol{s}}|\,\frac{\varepsilon R}{6R-\varepsilon-\varepsilon R} \leq \frac{\varepsilon}{4}\,|\Sigma_{\boldsymbol{r},\boldsymbol{s}}|,\end{aligned}$$

が成り立つ。よって

$$||\Sigma_{\boldsymbol{r},\boldsymbol{s}}| - \mathrm{E}[Z]| \leq \frac{\varepsilon}{4}\,|\Sigma_{\boldsymbol{r},\boldsymbol{s}}| + |\Sigma_{\boldsymbol{r},\boldsymbol{s}}|\prod_i(1-\hat{\rho}_i)^{M+1},$$

を得る。最後に $\prod_i(1-\hat{\rho}_i)^{M+1}$ の大きさを

$$\prod_i(1-\hat{\rho}_i)^{M+1} \leq \left\{1-\left(\frac{1}{m}-\frac{\varepsilon}{6mR}\right)\right\}^{RM} = \left\{1-\left(\frac{1-\frac{\varepsilon}{6R}}{m}\right)\right\}^{RM}$$

$$\begin{aligned} &\leq \left(1-\frac{5}{6m}\right)^{RM} \leq \left(1-\frac{5}{6m}\right)^{R\,108mR^2\varepsilon^{-2}\ln(2R/\delta)} \\ &\leq \left(\mathrm{e}^{-1}\right)^{\frac{5}{6m}108mR^3\varepsilon^{-2}\ln(2R/\delta)} \\ &= \mathrm{e}^{-90R^3\varepsilon^{-2}\ln(2R/\delta)}. \end{aligned}$$

従って、偏りの大きさは以下のようになる。

$$\begin{aligned} \left||\Sigma_{\boldsymbol{r},\boldsymbol{s}}| - \mathrm{E}[Z]\right| &\leq \frac{\varepsilon}{4}|\Sigma_{\boldsymbol{r},\boldsymbol{s}}| + \mathrm{e}^{-90R^3\varepsilon^{-2}\ln(2R/\delta)}|\Sigma_{\boldsymbol{r},\boldsymbol{s}}| \\ &\leq |\Sigma_{\boldsymbol{r},\boldsymbol{s}}|\left(\frac{\varepsilon}{4}+\mathrm{e}^{-90R^3\varepsilon^{-2}\ln(2R/\delta)}\right). \qquad \square \end{aligned}$$

この結果から、提案したアルゴリズムによって得られる近似解と厳密解との偏りの大きさは、主に標本分布と一様分布とのずれに依存し、モンテカルロ法を採用したことによる偏りはほとんど無いと言える。

5 結論と課題

我々のスキームは、$m \times n$ 分割表数え上げ問題に対して多項式回の問題縮小と多項式個の標本抽出で誤差の大きさと偏りの幅を確率論的に押さえられた近似解を与える。2002 年 Cryan et al. が行数が定数の 2 元分割表に対する heat bath マルコフ連鎖が rapid mixing であることを示した [1]。もちろんこれを我々のスキームに適用すると、我々のスキームも行数固定の $m \times n$ 分割表数え上げ問題に対する多項式時間確率論的近似解法となる。しかし、一般の $m \times n$ 分割表に対する多項式時間近似一様標本抽出法の存在性は現時点では未解決である。

参考文献

[1] M. Cryan, M. Dyer, L. A. Goldberg, M. Jerrum, and R. Martin, "Rapidly mixing Markov chains for sampling contingency tables with constant number of rows," *Proceedings of the 43rd Annual Symposium on Foundations of Computer Science (FOCS)* (2002), pp. 711–720.

[2] M. Dyer, A. Frieze, and R. Kannan, "A random polynomial time algorithm for approximating the volume of convex bodies," *Journal of the ACM*, 38 (1991), pp. 1–17.

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

[4] M. Dyer, R. Kannan, and J. Mount, "Sampling contingency tables," *Random Structures and Algorithms*, 10 (1997), pp. 487–506.

[5] D. Hernek, "Random generation of $2 \times n$ contingency tables," *Random Struct. Algorithms*, 13(1998), pp. 71–79.

[6] W. Höffding, "Probability inequalities for sums of bounded random variables," *J. Amer. Stat. Assoc.* 58(1963), pp. 13–30.

[7] S. Kijima and T. Matsui, "Approximate counting scheme for $m \times n$ contingency tables ," Tech. Rep., The Univ. of Tokyo, *METR* 03-01, January (2003).