УДК 519.2

Системы различных представителей для случайных множеств

В. Н. Сачков (Москва)

Введение

Обозначим через 2^X совокупность подмножеств конечного множества X, содержащего m элементов. Известно, что элементы $(x_1, x_2, \ldots, x_n)$ образуют систему различных представителей непустых множеств $X_1, X_2, \ldots, X_n$, где $X_i \in 2^X$, $i=1, 2, \ldots, n$, если $x_i \in X_i$, $i=1, 2, \ldots, n$, и $x_i \neq x_j$, $i \neq j$ (см. [1]).

Если $A=\|a_{ij}\|$, $i=1, 2, \ldots, n$, $j=1, 2, \ldots, m$, есть матрица инцидентности $X_1, X_2, \ldots, X_n$ и X, т. е. (0, 1)-матрица такая, что

$$a_{ij}=\begin{cases}1, & x_j \in X_i,\\ 0, & x_j \notin X_i,\end{cases}$$

то необходимое и достаточное условие существования системы различных представителей состоит в том, что

$$\operatorname{per} A > 0,$$

где $\operatorname{per} A$ — перманент матрицы A, определяемый равенством

$$\operatorname{per} A=\sum_{(j_1, j_2, \ldots, j_n)} a_{1j_1}a_{2j_2} \ldots a_{nj_n}.$$

Суммирование по $(j_1, j_2, \ldots, j_n)$ означает суммирование по всевозможным инъективным отображениям множества $Z_n=\{1, 2, \ldots, n\}$ в множество $Z_m=\{1, 2, \ldots, m\}$.

На совокупности непустых множеств, входящих в 2^X, зададим равномерное вероятностное распределение и будем предполагать, что множества выбираются независимо из 2^X. Возникает вопрос, с какой вероятностью множества $X_1, X_2, \ldots, X_n$, полученные по указанной вероятностной схеме, обладают системой различных представителей. В настоящей работе дается оценка такой вероятности и получается ее асимптотическое представление при $n, m \to \infty$. Из этого асимптотического представления следует основной результат данной статьи, состоящий в том, что при больших n и m с вероятностью, близкой к единице, независимо и равновероятно выбранные множества $X_1, X_2, \ldots, X_n$ имеют систему различных представителей.

Данный факт является полезным по следующей причине. Известно, что проверка необходимых и достаточных условий Ф. Холла для сущестбования системы различных представителей сводится к проверке 2^n-1 неравенств для чисел элементов в объединениях подмножеств множества из m элементов и при больших n и m весьма громоздка. По этой причине алгоритмы для отыскания систем различных представителей приходится применять при отсутствии информации об успешном решении задачи. Сформулированный выше основной результат данной статьи позволяет утверждать, что при больших n и m система различных представителей всегда существует и может быть найдена, например, с помощью известного алгоритма М. Холла [2].

Следует отметить, что рассматриваемый круг вопросов тесно связан с изучением асимптотических условий равенства нулю перманента случайной (0,1)-матрицы $n\times m$, когда n и $m\to\infty$. Эти вопросы рассматривались, в частности, в работе [3].

§ 1. Точные формулы и оценки

Будем считать, что $n\leqslant m$ и матрица A из множества $\mathfrak{A}_{nm}$ всех (0,1)-матриц $n\times m$ принадлежит:

а) множеству $\mathfrak{A}^0_{nm}$, если она не имеет нулевых строк;

б) множеству $\mathfrak{A}^{00}_{nm}$, если она не имеет нулевых линий;

в) множеству $\mathfrak{A}^1_{nm}$, если она имеет по крайней мере одну нулевую линию;

г) множеству $\overline{\mathfrak{A}}_{nm}$ ($\overline{\overline{\mathfrak{A}}}_{nm}$), если она имеет ровно одну нулевую строку (столбец), при этом

$$\mathfrak{A}^*_{nm}=\overline{\mathfrak{A}}_{nm}\cup\overline{\overline{\mathfrak{A}}}_{nm}.$$

Полагая, что $|Z|$ — число элементов конечного множества Z, введем обозначения

$$C_{nm}=|\mathfrak{A}^{00}_{nm}|,\quad Q_{nm}=|\mathfrak{A}^1_{nm}|,\quad Q^*_{nm}=|\mathfrak{A}^*_{nm}|.$$

Нетрудно видеть, что

$$|\mathfrak{A}^0_{nm}|=(2^m-1)^n,\quad C_{nm}+Q_{nm}=2^{nm},$$
$$Q^*_{nm}=n(2^m-1)^{n-1}+m(2^n-1)^{m-1}-nmC_{n-1,m-1}. \tag{1}$$

Обращая очевидные соотношения

$$\sum_{j=0}^{m-1}\binom{m}{j}C_{n,m-j}=(2^m-1)^n,\quad \sum_{j=0}^{n-1}\binom{n}{j}C_{n-j,m}=(2^n-1)^m$$

или используя метод включения — исключения, можно получить формулы

$$C_{nm}=\sum_{j=0}^{m}(-1)^j\binom{m}{j}(2^{m-j}-1)^n,\quad C_{nm}=\sum_{j=0}^{n}(-1)^j\binom{n}{j}(2^{n-j}-1)^m. \tag{2}$$

Введем еще одно обозначение:

$$N_{nm}=|\{A:\ \operatorname{per}A=0,\ A\in\mathfrak{A}^{00}_{nm}\}|.$$

Согласно теореме Кенига [1] равенство нулю перманента неотрицательной матрицы A размеров $n\times m$, $n\leqslant m$, эквивалентно наличию в A нулевой подматрицы $s\times r$, где $s+r>m$. Не ограничивая общности, считаем, что если $A\in\mathfrak{A}_{nm}^{00}$ и $\operatorname{per}A=0$, то матрица A имеет вид

$$A=\left(\begin{array}{c|c} A_1 & A_3 \\ \hline 0 & A_2 \end{array}\right),$$

где в левом нижнем углу расположена нулевая подматрица размера $r\times s$, а подматрицы A_1 и A_2 не содержат нулевых строк и столбцов. Числа r и s выбраны так, что сумма $r+s$ максимальна. Из этих соображений следует, что

$$N_{nm}\leqslant\sum_{r=2}^{n-1}\sum_{s=m-r+1}^{m-1}\binom{n}{r}\binom{m}{s}2^{(n-r)(m-s)}C_{n-r,s}C_{r,m-s}.$$

Так как

$$C_{nm}\leqslant(2^m-1)^n\leqslant 2^{mn},$$

то после очевидных упрощений имеем

$$N_{nm}\leqslant 2^{nm}\sum_{r=2}^{n-1}\sum_{s=m-r+1}^{m-1}\binom{n}{r}\binom{m}{s}\frac{1}{2^{rs}}. \tag{3}$$

Перейдем теперь к выводу формулы для вероятности того, что перманент случайной (0,1)-матрицы положителен. Обозначим через $E(j)$ событие, состоящее в том, что случайная равновероятная матрица $A\in\mathfrak{A}_{nm}^{0}$ имеет ровно j нулевых столбцов. Положим

$$E_{nm}=\bigcup_{j=m-n+1}^{m-1}E(j).$$

Для вероятности события $(\operatorname{per}A=0)$, состоящего в том, что перманент случайной матрицы $A\in\mathfrak{A}_{nm}^{0}$ равен нулю, имеет место формула

$$\mathbf{P}_{nm}(\operatorname{per}A=0)=\mathbf{P}(E_{nm})+\sum_{j=0}^{m-n}\mathbf{P}(E(j)\cap(\operatorname{per}A=0)).$$

Очевидно, что

$$\mathbf{P}(E(j))=\binom{m}{j}\frac{C_{n,m-j}}{(2^m-1)^n},\quad j=0,1,\ldots,m-1;$$

$$\mathbf{P}(E(j)\cap(\operatorname{per}A=0))=\binom{m}{j}\frac{N_{n,m-j}}{(2^m-1)^n},\quad j=0,1,\ldots,m-1.$$

Используя эти равенства, получаем формулу для вероятности того, что перманент случайной матрицы $A\in\mathfrak{A}_{nm}^{0}$ положителен:

$$\mathbf{P}_{nm}(\operatorname{per}A>0)=\frac{1}{(2^m-1)^n}\sum_{j=0}^{m-n}\binom{m}{j}(C_{n,m-j}-N_{n,m-j}). \tag{4}$$

Полагая

$$\mathbf{P}_n(\operatorname{per} A>0)=\mathbf{P}_{nn}(\operatorname{per} A>0), \quad C_n=C_{n,n}, \quad N_n=N_{n,n},$$

отсюда получаем

$$\mathbf{P}_n(\operatorname{per} A>0)=\frac{C_n-N_n}{(2^n-1)^n}. \tag{5}$$

Соответствующая вероятность положительности перманента для $A\in\mathfrak{A}_{nn}$ равна

$$\widetilde{\mathbf{P}}_n(\operatorname{per} A>0)=\frac{C_n-N_n}{2^{n^2}}. \tag{6}$$

Из формул (2)—(4) получаем оценку

$$\mathbf{P}_{nm}(\operatorname{per} A>0)\geqslant\frac{1}{(2^m-1)^n}\sum_{j=0}^{m-n}\binom{m}{j}\left[\sum_{l=0}^{m-j}(-1)^l\binom{m-j}{l}(2^{m-j-l}-1)^n-\right.$$
$$\left.-2^{nm}\sum_{r=2}^{n-1}\sum_{s=m-r+1}^{m-1}\binom{n}{r}\binom{m}{s}\frac{1}{2^{rs}}\right]. \tag{7}$$

При $m=n$ это неравенство можно упростить следующим образом:

$$\mathbf{P}_n(\operatorname{per} A>0)\geqslant\Phi_n=\frac{1}{2^{n^2}}\sum_{j=0}^{n}(-1)^j\binom{n}{j}(2^{n-j}-1)^n-\sum_{r=2}^{n-1}\sum_{s=n-r+1}^{n-1}\binom{n}{r}\binom{n}{s}\frac{1}{2^{rs}}. \tag{8}$$

Значения Φ_n при $5\leqslant n\leqslant 20$ указаны в таблице:

n	Φ_n	n	Φ_n
5	0,134262956	13	0,996706783
6	0,501603881	14	0,998254045
7	0,759970121	15	0,999073192
8	0,892653562	16	0,999508269
9	0,951065137	17	0,999739490
10	0,976353600	18	0,999862380
11	0,988010399	19	0,999927401
12	0,993755795	20	0,999961838

Из таблицы видно, что уже при сравнительно небольших значениях n оценка Φ_n близка к единице.

§ 2. Предельные теоремы

Основная цель настоящего параграфа — показать, что при $n\leqslant m$ для больших m случайно выбранные непустые множества $X_1, X_2, \ldots, X_n\in 2^X$, $|X|=m$, почти всегда обладают системой различных представителей. Если при $n\leqslant m$ имеем дополнительное условие, состоящее в том, что $n\to\infty$, то справедлив даже более содержательный результат.

Теорема 1. *Пусть* $n \leqslant m$ *и* U_{nm} — *событие, состоящее в том, что для любого фиксированного набора различных элементов* $x_1, x_2, \ldots, x_n$ *найдется такая подстановка* s *степени* n, *что система элементов* $(x_{s(1)}, x_{s(2)}, \ldots, x_{s(n)})$ *будет системой различных представителей для случайно выбранных непустых множеств* $X_1, X_2, \ldots, X_n \in 2^X$ *соответственно, причем* $|X| = m$. *Тогда для вероятности* $\mathbf{P}(U_{nm})$ *того, что событие* U_{nm} *осуществится, имеет место равенство*

$$\lim_{n\to\infty} \mathbf{P}(U_{nm}) = 1.$$

Пусть A' — подматрица матрицы A, являющаяся матрицей инцидентности множеств $X_1, X_2, \ldots, X_n$ и множества $X' = \{x_1, x_2, \ldots, x_n\}$. Ясно, что

$$\mathbf{P}_{nm}(\operatorname{per} A > 0) \geqslant \widetilde{\mathbf{P}}_n(\operatorname{per} A' > 0).$$

Поэтому для доказательства теоремы достаточно показать, что

$$\lim_{n\to\infty} \widetilde{\mathbf{P}}_n(\operatorname{per} A' > 0) = 1. \tag{9}$$

Полагая $m = n$, из формулы (3) получаем

$$\frac{N_n}{2^{n^2}} \leqslant \sum_{r=2}^{n-1} \sum_{s=n-r+1}^{n-1} \binom{n}{r}\binom{n}{s}\frac{1}{2^{rs}}. \tag{10}$$

При $n \geqslant 3$ имеют место следующие оценки:

$$\sum_{s=2}^{n-1}\binom{n}{s}\frac{1}{2^{(n-1)s}} \leqslant \frac{n(n-1)}{2^{2n-1}} + \frac{1}{2^{3(n-1)}}\sum_{s=3}^{n-1}\binom{n}{s} \leqslant \frac{n(n-1)+4}{2^{2n-1}},$$

$$\sum_{r=3}^{n-2}\sum_{s=n-r+1}^{n-1}\binom{n}{r}\binom{n}{s}\frac{1}{2^{rs}} \leqslant 2^n \sum_{r=3}^{n-2}\binom{n}{r}\frac{1}{2^{r(n-r+1)}} \leqslant$$

$$\leqslant \frac{1}{2^{2n-6}}\sum_{r=3}^{[(n+1)/2]}\binom{n+1}{r} \leqslant \frac{1}{2^{n-6}}.$$

Используя полученные оценки и неравенство (10), имеем

$$\frac{N_n}{2^{n^2}} \leqslant \frac{n^2(n-1)}{2^{2n-1}} + \frac{n(n-1)+4}{2^{2n-1}} + \frac{1}{2^{n-6}}.$$

Отсюда при $n \geqslant 3$ следует, что

$$\frac{N_n}{2^{n^2}} \leqslant \frac{9}{2^{n-3}}. \tag{11}$$

Из формул (2) при $m = n$ и неравенства Бонферрони следует, что

$$(2^n - 1)^n - n(2^{n-1} - 1)^n \leqslant C_n \leqslant (2^n - 1)^n.$$

Отсюда, принимая во внимание неравенство

$$(1 - z)^n \geqslant 1 - nz, \quad 0 \leqslant z \leqslant 1, \quad n \geqslant 1,$$

находим

$$1-\frac{n}{2^{n-1}}\leqslant\frac{C_n}{2^{n^2}}\leqslant 1. \tag{12}$$

Из равенства (6), используя оценки (11) и (12), окончательно получаем

$$\widetilde{\mathbf{P}}_n(\operatorname{per} A'>0)\geqslant 1-\frac{n}{2^{n-1}}-\frac{9}{2^{n-3}}.$$

Отсюда следует равенство (9) и, следовательно, справедливость теоремы 1.

Следствие. *Если $n\leqslant m$ и $n\to\infty$, то с вероятностью, стремящейся к единице, случайный набор непустых множеств $X_1, X_2, \ldots, X_n \in 2^X$, $|X|=m$, обладает системой различных представителей.*

Сформулируем и докажем теперь общую теорему.

Теорема 2. *При $n\leqslant m$ и $m\to\infty$ с вероятностью, стремящейся к единице, случайный набор непустых множеств $X_1, X_2, \ldots, X_n \in 2^X$, $|X|=m$, обладает системой различных представителей.*

Если $n\to\infty$, то данная теорема совпадает со следствием из теоремы 1. Поэтому остается рассмотреть случай, когда $n=\mathrm{const}$, $m\to\infty$.

Применяя ко второй из формул (2) неравенство Бонферрони, имеем:

$$(2^n-1)^{m-j}-n(2^{n-1}-1)^{m-j}\leqslant C_{n,m-j}\leqslant(2^n-1)^{m-j}.$$

Суммируя эти неравенства по $j=0, 1, \ldots, m-n$ и деля на 2^{nm}, получаем при $n>1$:

$$1-\frac{n}{2^m}-\frac{1}{2^{(n-1)(m-n)}}\leqslant\frac{1}{2^{nm}}\sum_{j=0}^{m-n}\binom{m}{j}C_{n,m-j}\leqslant 1. \tag{13}$$

Заметим, что при $n=1$

$$\frac{1}{2^m}\sum_{j=0}^{m-1}\binom{m}{j}C_{1,m-j}=1-\frac{1}{2^m}.$$

Правая часть неравенства (3) отлична от нуля только при $n\geqslant 3$, поэтому для таких n имеем

$$\frac{N_{mn}}{2^{nm}}\leqslant\sum_{s=m-n+2}^{m-1}\binom{m}{s}\sum_{r=m-s+1}^{n-1}\binom{n}{r}\frac{1}{2^{rs}}\leqslant\frac{2^{(n-1)(n-2)+n}}{2^{m(n-2)}}.$$

Отсюда следует оценка при $n\geqslant 3$:

$$\frac{1}{2^{nm}}\sum_{j=0}^{m-n}\binom{m}{j}N_{n,m-j}\leqslant 2^{(n-1)(n-2)+n}\left(\frac{5}{2^n}\right)^m. \tag{14}$$

Из равенства (4), используя при $n\geqslant 3$ оценки (13) и (14), находим:

$$\mathbf{P}_{nm}(\operatorname{per} A>0)\geqslant 1-\frac{n}{2^m}-\frac{1}{2^{(n-1)(m-n)}}-2^{(n-1)(n-2)+n}\left(1-\frac{1}{2^m}\right)^{-n}\left(\frac{5}{2^n}\right)^m. \tag{15}$$

Так как $N_{1,m}=N_{2,m}=0$, то ввиду формулы (4)

$$\mathbf{P}_{1m}(\operatorname{per} A > 0) = 1, \quad \mathbf{P}_{2m}(\operatorname{per} A > 0) = 1 - \frac{mC_{2,2}}{(2^m-1)^2}. \tag{16}$$

Из неравенства (15) и равенств (16) следует, что при $n=\text{const}$

$$\lim_{m\to\infty} \mathbf{P}_{nm}(\operatorname{per} A > 0) = 1.$$

Этим теорема 2 доказана.

Теорема 3. *Пусть $n \leqslant m$ и U_{nm}— событие, состоящее в том, что для любого фиксированного набора $x_1, x_2, \ldots, x_n$ различных элементов из X, $|X|=m$, существует подстановка s степени n, при которой $(x_{s(1)}, x_{s(2)}, \ldots \ldots, x_{s(n)})$ является системой различных представителей случайных множеств $X_1, X_2, \ldots, X_n \in 2^X$ и пусть для того же набора $x_1, x_2, \ldots, x_n$ определены события V'_{nm}, V''_{nm} и V_{nm}: V'_{nm} состоит в том, что какой-то один элемент из совокупности $x_1, x_2, \ldots, x_n$ не принадлежит ни одному из случайных множеств $X_1, X_2, \ldots, X_n$; V''_{nm} заключается в том, что какое-то одно из случайных множеств $X_1, X_2, \ldots, X_n$ не содержит ни одного из элементов $x_1, x_2, \ldots, x_n$;*

$$V_{nm} = \overline{V}'_{nm} \cap \overline{V}''_{nm} = \overline{V'_{nm} \cup V''_{nm}}.$$

Тогда при $n\to\infty$ события U_{nm} и V_{nm} асимптотически эквивалентны, т. е.

$$\lim_{n\to\infty} \mathbf{P}(U_{nm} \mid V_{nm}) = \lim_{n\to\infty} \mathbf{P}(V_{nm} \mid U_{nm}) = 1.$$

Если A'— матрица инцидентности $X'=\{x_1, x_2, \ldots, x_n\}$ и случайных множеств $X_1, X_2, \ldots, X_n$, то для доказательства теоремы достаточно показать, что

$$\lim_{n\to\infty} \mathbf{P}(A' \in \mathfrak{A}^*_{nn} \mid \operatorname{per} A' = 0) = 1.$$

Нетрудно получить следующее выражение для интересующей нас вероятности:

$$\mathbf{P}(A' \in \mathfrak{A}^*_{nn} \mid \operatorname{per} A'=0) = \frac{\mathbf{P}(A' \in \mathfrak{A}^*_{nn})}{\mathbf{P}(A' \in \mathfrak{A}^1_{nn}) + \mathbf{P}(A' \in \mathfrak{A}^{00}_{nn} \cap (\operatorname{per} A' = 0))}.$$

Отсюда следует, что

$$\mathbf{P}(A' \in \mathfrak{A}^*_{nn} \mid \operatorname{per} A' = 0) = \frac{Q^*_n}{Q_n + N_n}, \tag{17}$$

где $Q_n = Q_{nn}$, $Q^*_n = Q^*_{nn}$.

Нетрудно убедиться в том, что

$$\lim_{n\to\infty} \frac{nC_{n-1}}{2^{n^2-n+1}} = 0.$$

Поэтому из формулы (1) при $m=n$ следует соотношение

$$\lim_{n\to\infty} \frac{Q_n^*}{n2^{n^2-n+1}} = 1.$$

Кроме того, используя формулу (2) при $m=n$ и соотношение $C_n+Q_n=$ $=2^{n^2}$, можно получить равенство

$$\lim_{n\to\infty} \frac{Q_n}{n2^{n^2-n+1}} = 1.$$

Наконец, из неравенства (11) следует, что

$$\lim_{n\to\infty} \frac{N_n}{n2^{n^2-n+1}} = 0.$$

Теперь справедливость теоремы следует из формулы (17).

(Поступила в редакцию 24/I 1975 г.)

Литература

1. **Г. Дж. Райзер,** Комбинаторная математика, Москва, изд-во «Мир», 1966.
2. **М. Холл,** Комбинаторика, Москва, изд-во «Мир», 1970.
3. **C. I. Everett, P. R. Stein,** The asymptotic number of (0,1)-matrices with zero permanent, Discrete Math., **6** (1973), 29—34.